1. Sets
    1. Sets
    2. Subsets
    3. Union of sets
    4. Intersection of sets
    5. Difference of sets
    6. Working with two sets
    7. Working with three sets
    8. Cartesian product
    9. Cartesian Product (exercise)
2. Relations
    1. Relations
    2. Relation types
    3. Functions
    4. Function types
    5. Composition of functions
    6. Inverse of a function
3. Real functions
    1. Domain and range
    2. Inequality
    3. Polynomial function
    4. Quadratic polynomial function
    5. Rational function
    6. Rational inequality
    7. Function operations
    8. Modulus function
    9. Greatest and least integer functions
    10. Exponential and logarithmic functions
    11. Domain and range of exponential and logarithmic function
    12. Trigonometric functions
    13. Trigonometric values, equations and identities

    14. Trigonometric inequalities
    15. Inverse trigonometric functions
    16. Composition of trigonometric function and its inverse
    17. Function operations (exercise)
    18. Composition of functions (exercise)
    19. Value of a function
4. Transformation of graphs
    1. Transformation of graphs
    2. Transformation of graph by output
    3. Transformation of graphs by modulus function
    4. Transformation of graphs by greatest integer function
    5. Transformation applied by fraction part function
5. Function properties
    1. Equal functions
    2. Even and odd functions
    3. Periodic functions
    4. Periodicity and periods
    5. Monotonic functions
    6. Increasing and decreasing intervals
    7. Minimum and maximum values
    8. Least and greatest function values
    9. One-one and many-one functions (exercise)
    10. Inverse functions (exercise)
    11. Limits
    12. Limits of algebraic functions
    13. Continuous function

Sets

Set theory is about studying collection of objects. The collection may comprise anything or any abstraction. It can be purely abstract thing like numbers or abstraction of real thing like students studying in class XI in a school. The members of collection can be numbers, letters, titles of books, people, teachers, provinces – virtually anything - even other collections. Further, it need not be finite. For example, a set of integers has infinite members. For a set, only requirement is that the members of a collection are properly defined.

Set

> A set is a collection of well defined objects.

In other words, the member of set is clearly identifiable. The terms “object”, “member” or “element” mean same thing and are used interchangeably.

How to represent a set?

A set is denoted by capital letters like “A”, “B”, “C” etc. In choosing a symbol for a set, it is generally convenient to use a capital letter that identifies with the set. For example, it is appropriate to use symbol “V” to represent collection of vowels in English alphabet.

On the other hand, the members or elements of a set are denoted by small letters like “a”,”b”,”c” etc.

Membership of a set is denoted by the symbol “ $\in$” . Its literal meaning is “belongs to”. If an object does not belong to a set, then we convey the same, using symbol “ $\notin$ ”.

$a \in A$: we read this as “a” belongs to set "A".

$a \notin A$: we read this as “a” does not belong to set "A".

The set is represented in two ways :

- Roaster form
- Set builder form

Roaster form

All elements of the set are listed with a comma (“,”) in between and the listing itself is enclosed within braces “{“ and “}”. The order or sequence of elements within the set is not important – though desirable.

The set of numbers, which divide 12, is written as :

$$A = \{1,2,3,4,6,12\}$$

If a pattern or sequence is easily made out, then we can use ellipsis ("...") to represent continuity of such pattern. This type of representation is particularly useful to represent an infinite set. Clearly, sequence of members in this type of representation is important.

The set of even numbers is written as,

$$B = \{2,4,6,8\ldots\ldots\ldots\}$$

The roaster form is limited in certain circumstance. For example, we can not represent set of real numbers in roaster form. Real numbers is an infinite set, but the elements of this set do not follow a pattern or have a particular sequence. As such, we can not define same with the help of ellipsis.

Every member of the set is unique and distinct. However, we encounter situations in which collection can have repeated elements. For example, the set representing scores of three students can be a set of three numbers one of which is repeated :

$$S = \{80,80,70\}$$

We need to reduce such collection as :

$$\Rightarrow S = \{80,80,70\} = \{80,70\}$$

Set builder form

Collections are often characterized by a common property. We can, therefore, define members of a set in terms of the common property. However, we need to ensure that objects outside the collection do not have the same common property.

The construction of qualification for the common property is quite flexible. Only thing is that we need to be explicit in what we mean. Generally, we denote the member by a symbol like “x” and then define the membership. Consider the examples :

$$A = \{\text{x: x is a vowel in English alphabet}\}$$

$$B = \{\text{x: x is an integer and } 0 < x < 10\}$$

The roaster equivalents of two sets are :

$$A = \{a, e, i, o, u\}$$

$$B = \{1,2,3,4,5,6,7,8,9\}$$

Can we write set “B” as the one which comprises single digit natural number? Yes. Thus, we can see that there are indeed different ways to define and identify members and hence the flexibility in defining collection.

We should be careful in using words like “and” and “or” in writing qualification for the set. Consider the example here :

$$C = \{\text{x:} x \in Z \text{ and } 2 < x < 4\}$$

Both conditional qualifications are used to determine the collection. The elements of the set as defined above are integers. Thus, the only member of the set is “3”.

Now, let us consider an example, which involves “or” in the qualification,

$$C = \{\text{x: } x \in A \text{ or } x \in B\}$$

The member of this set can be elements belonging to either of two sets "A" and "B". The set consists of elements (i) belonging exclusively to set "A", (ii) elements belonging exclusively to set "B" and (iii) elements common to sets "A" and "B".

Example

Problem 1 : A set in roaster form is given as :

$$A = \left\{ \frac{5^2}{6}, \frac{6^2}{7}, \frac{7^2}{8} \right\}$$

Write the set in “set builder form”.

Solution : We see here that we are dealing with natural numbers. The numerators are square of natural numbers in sequence. The number in denominator is one more than numerator for each member. We can denote natural number by “n”. Clearly, if numerator is “ n^2 ”, then denominator is “n+1”. Therefore, the expression that represent a member of the set is :

$$x = \frac{n^2}{n+1}$$

However, this set is not an infinite set. It has exactly three members. Therefore, we need to specify “n” so that only members of the set are exclusively denoted by the above expression. We see here that “n” is greater than 4, but “n” is less than 8. For representing three elements of the set,

$$5 \le n \le 7$$

We can write the set, now, in the builder form as :

$$A = \left\{ x : \quad x = \frac{n^2}{n+1}, \text{where "n" is a natural number and} \quad 5 \le n \le 7 \right\}$$

In set builder form, the sequence within the range is implied. It means that we start with the first valid natural number and proceed sequentially till the last valid natural number.

Some important sets representing numbers

Few key number sets are used regularly in mathematical context. As we use these sets often, it is convenient to have predefined symbols :

- P(prime numbers)

- N (natural numbers)
- Z (integers)
- Q (rational numbers)
- R (real numbers)

We put a superscript “+”, if we want to specify membership of only positive numbers, where appropriate. " Z^+ ", for example, means set of positive integers.

Empty set

An empty set has no member or object. It is denoted by symbol “φ” and is represented by a pair of braces without any member or object.

$$\varphi = \{\}$$

The empty set is also called “null” or “void” set. For example, consider a definition : “the set of integer between 1 and 2”. There is no integer within this range. Hence, the set corresponding to this definition is an empty set. Consider another example :

$$B = \left\{x : \quad x^2 = 4 \quad \text{and x is odd}\right\}$$

An odd integer squared can not be even. Hence, set “B” also does not have any element in it.

There is a bit of paradox here. If the definition does not yield an element, then the set is not well defined. We may be tempted to say that empty set is not a set in the first place. However, there is a practical reason to have an empty set. It enables mathematical operations. We shall find many examples as we study operations on sets.

Equal sets

The members of two equal sets are exactly same. There is nothing more to it. However, we need to know two special aspects of this equality. We mentioned about repetition of elements in a set. We observed that repetition of elements does not change the set. Consider example here :

$$A = \{1{,}5{,}5{,}8{,}7\} = \{1{,}5{,}8{,}7\}$$

Another point is that sequence does not change the set. Therefore,

$$A = \{1{,}5{,}8{,}7\} = \{5{,}7{,}8{,}1\}$$

In the nutshell, when we have to compare two sets we look for distinct elements only. If they are same, then two sets in question are equal.

Cardinality

Cardinality is the numbers of elements in a set. It is denoted by modulus of set like |A|.

Cardinality

> The cardinality of a set “A” is equal to numbers of elements in the set.

The cardinality of an empty set is zero. The cardinality of a finite set is some positive integers. The cardinality of a number system like integers is infinity. Curiously, the cardinality of some infinite set can be compared. For example, the cardinality of natural numbers is less than that of integers. However, we can not make such deduction for the most case of infinite sets.

Subsets

The collections are generally linked in a given context. If we think of ourselves, then we belong to a certain society, which in turn belongs to a province, which in turn belongs to a country and so on. In the context of a school, all students of a school belong to school. Some of them belong to a certain class. If there are sections within a class, then some of these belong to a section.

We need to have a mathematical relationship between different collections of similar types. In set theory, we denote this relationship with the concept of “subset”.

Subset

> A set, “A” is a subset of set “B”, if each member of set “A” is also a member of set “B”.

We use symbol “ $\subset$ ” to denote this relationship between a “subset” and a “set”. Hence,

$$A \subset B$$

We read this symbolic representation as : set “A” is a subset of set “B”. We express the intent of relationship as :

$$A \subset B \quad \text{if} \quad x \in A, \quad \text{then} \quad x \in B$$

It is evident that set "B" is larger of the two sets. This is sometimes emphasized by calling set "B" as the “superset” of "A". We use the symbol “ $\supset$ ” to denote this relation :

$$B \supset A$$

If set "A" is not a subset of "B", then we write this symbolically as :

$$A \not\subset B$$

Important results / deductions

Some of the important characteristics and related deductions are presented here :

Equality of two sets

It is clear that set “B” is inclusive of subset “A”. It means that “B” may have additional elements over and above those common with “B”. In case, all elements of “B” are also in “A”, then two sets are equal. We express this symbolically as :

$$If \quad A \subset B \quad \text{and} \quad B \subset A, \quad \text{then} \quad A = B.$$

This is true in other direction as well :

$$If \quad A = B, \quad \text{then} \quad A \subset B \quad \text{and} \quad B \subset A.$$

We can write two instances in a single representation as :

$$A \subset B \quad \text{and} \quad B \subset A \Leftrightarrow A = B$$

The symbol “ $\Leftrightarrow$ ” means that relation holds in either direction.

Relation with itself

Every set is subset of itself. This is so because every element is present in itself.

Relation with Empty set

Empty set is a subset of every set. This deduction is direct consequence of the fact that empty set has no element. As such, this set is subset of all sets.

Proper subset

We have seen from the deductions above that special circumstance of “equality” can blur the distinction between “set” and “subset”. In order to emphasize, mother-child relation between sets, we coin the term “proper subset”. If every element of set “B” is not present in set “A”, then “A” is a “proper” subset of set “B”; otherwise not. This means that set “B” is a larger set, which, besides other elements, also includes all elements of set “A”.

Set of vowels in English alphabet, “V”, is a “proper” subset of set of English alphabet, “E”. All elements of “V” are present in “E”, but not all elements of “E” are present in “V”.

There is a bit of conventional differences. Some write a “proper” subset relation using symbol “ $\subset$ ” and write symbol “ $\subseteq$ ” to mean possibility of equality as well. We have chosen not to differentiate two subset types.

Number system

The number system is one such system, in which different number groups are related. Natural number is a subset of integers. integers are subset of rational numbers and rational numbers are subset of real numbers. None of these sets are equal. Hence, relations are described by proper subsets.

$$N \subset Z$$

We can write the chain of relation among number sets :

$$\Rightarrow P \subset N \subset Z \subset Q \subset R$$

However, irrational numbers are also subset of real numbers, but irrational numbers is not rational numbers. We represent this relation by emphasizing that rational numbers is not a subset of irrational numbers or vice-versa. We depict this relation as :

$$Q(\text{rational numbers}) \not\subset T(\text{irrational numbers})$$

But irrational numbers is subset of real numbers. The real numbers comprises of only two subsets at the highest level – rational and irrational.

Therefore, irrational numbers is the remaining collection after deducting rational numbers from real numbers.

Following the logic, we define set of irrational numbers as :

$$T(\text{irrational numbers}) = \{x : x \in R \quad \text{and} \quad x \notin Q\}$$

Power set

Power set is formed of all possible subsets of a given set. It is denoted as P(A).

Power set

> The collection of all subsets of a set “A” is called power set, P(A).

For example, consider a set given by :

$$A = \{1,3,4\}$$

What are the possible subsets? There are three subsets consisting of individual elements: {1}, {3} and {4}. Then, elements taken two at a time form following subsets : {1,3}, {1,4} and {3,4}. Since order or sequence does not matter in set representation, there are only three subsets of two elements taken together. Now, the elements taken three at a time form the only one subset : {1,3,4}. Remember, a set is a subset of itself. Further, empty set (φ) is subset of any set. Hence, φ is also a subset of the given set “A”.

The set comprising of all possible subsets of given set “A” is :

$$P(A) = \{\varphi, \{1\}, \{3\}, \{4\}, \{1,3\}, \{1,4\}, \{3,4\}, \{1,3,4\}\}$$

We note two important points from this representation of power set :

1: The elements of a power set are themselves sets. In other words, every element of a power set is a set.

2: If the numbers of elements (cardinality) in a set is “n”, then numbers of elements in power set is 2^n .

For a set having three elements, the total numbers of elements in the power set is :

$$\Rightarrow m = 2^n = 2^3 = 8$$

We can see that this result is consistent with the illustration given above. We should, here, emphasize to avoid confusion that counting of elements of a set (cardinality) excludes empty set. It is, however, counted as members of power set.

Example

Problem 1: The finite sets “A” and “B” have “m” and “n” numbers of elements respectively. The total numbers of subsets of “A” is 56 more than the total numbers of subsets of “B”. Find “m” and “n”.

Solution : According to relation obtained for power set, the total numbers of subsets of “A” and “B” are :

$$k_A = 2^m$$

$$k_B = 2^n$$

According to question,

$$k_A - K_B = 56$$

$$\Rightarrow 2^m - 2^n = 56$$

We need to find two equations to find “m” and “n”. For this we seek expansion of “56” in terms of powers of “2”.

$$56 = 8X7 = 8(8-1) = 2^3\left(2^3 - 1\right)$$

In order to get this form, we rearrange the expression on the LHS of the earlier equation as :

$$\Rightarrow 2^n\left(2^{m-n}-1\right)=2^3\left(2^3-1\right)$$

Equating powers of similar base,

$$n=3 \quad \text{and} \quad m=6$$

Intervals

Intervals is an alternative way to represent a subset of real numbers. Real numbers is represented by a number line having infinite membership. We can think any segment of this number line as subset or interval. Consider an interval, where "a" and "b" belongs to real numbers and a < b :

$$a<x<b$$

The value of "x" falls between "a" and "b". For example, an interval 2<x<4 is a collection of all points lying between end points 2 and 4. The important thing is that this interval does not include end points and is called "open" interval. We can represent this collection as a set in "set builder form" as :

$$\{x : x \in R \quad and \quad 2<x<4\}$$

Alternatively, we can use pair of small brackets to represent open interval as :

$$(2,4)$$

The two forms of representations are equivalent. The later form is obviously an easier and convenient representation of the subset of real number. We use small braket "(" or ")" to denote interval that excludes end point. Likewise, we use square bracket "[" or "]" to denote interval that includes end point. We can represent a "close" interval as [2,4]. This interval is equivalent to :

$$[2,4]=2 \le x \le 4$$

We can have combination of "open" and "close" brackets like :

$$(2,4]=2<x \le 4$$

As a reminder, we should note that interval corresponding to real numbers or its subset is an infinite set as we can have infinite points on the line segment corresponding to an interval.

Graphical representation

The graphical representation uses a segment of line on the number line representing real numbers. The line segment is demarcated by a pair of two small circles – a filled circle to mean that end point is included in the interval and an unfilled circle to mean that end point is excluded from the interval.

Let us consider $a,b \in R$ and $a < b$, then

$$(a,b) = a < x < b$$

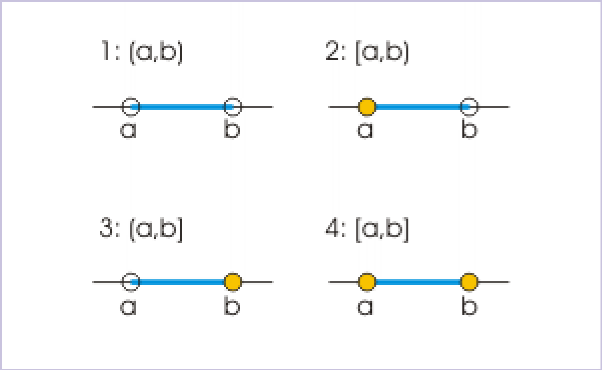
$$[a,b) = a \leq x < b$$

$$(a,b] = a < x \leq b$$

$$[a,b] = a \leq x \leq b$$

Graphically,

Intervals



Representation on real number line.

Set of real numbers

The real numbers is represented graphically by a straight line. The question that we seek to be answer here is whether the set of integers is bounded by infinity. In other words, whether we can define interval of real numbers like :

$$[-\infty, \infty]$$

The literal meaning of infinity is “unboundedness”. Infinity is considered as a large number, which may either be positive or negative. It does not have a finite (fixed) value. Infinity, therefore, is not a part of real number system. It does not lie on the real number line. For this reason, we can not assign infinity to a real variable like (though we do generally):

$$x = \infty$$

It follows, then, that appropriate interval, representing real numbers, is open at both ends :

$$R = -\infty < x < \infty = (-\infty, \infty)$$

Interval of real numbers greater than or less than a given value

In the interval form, we can write the set of real numbers greater than a given value, "a", as :

$$a < x < \infty = (a, \infty)$$

This is equivalent to :

$$x > a$$

The final notation "x > a" does not require to mention about infinity. It is an interval of real numbers greater than the given value 'a' appearing on the right. It is implied that it can be any large value. Similarly, the interval of real numbers less than a given value is :

$$-\infty < x < a = (-\infty, a)$$

$$x < a$$

Union of sets

We are familiar with basic algebraic operations. These basic mathematical operations, however, are not valid in all contexts. For example, algebraic operation such as addition has different details, when operated on vectors. Clearly, we expect that these operations will also be not same in the case of sets – which are collections and not individual elements.

Nevertheless, set operations bear resemblance to algebraic operation. For example, when we combine (not add) two sets, then the operation involved is called “union”. We can see that there is resemblance of the intent of addition, subtraction etc in the case of sets also.

Venn diagrams

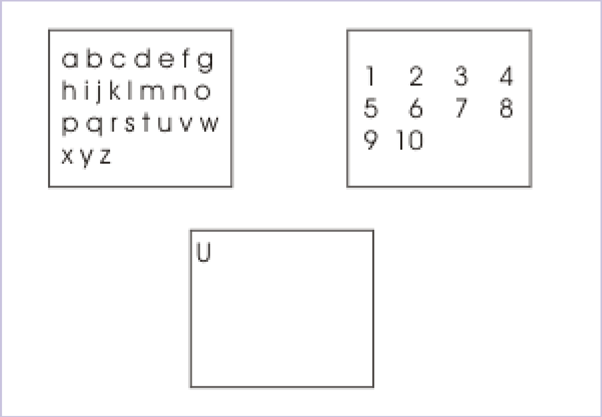
Venn diagrams are pictorial representation of sets/subsets and relationship that the sets/subsets have among them. It helps us to analyze relationship and carry out valid set operations in a relatively easier manner vis – a – vis symbolic representation.

Universal set

Universal set is the largest set among collection of sets. Importantly, it is not the collection of everything as might be conjectured by the nomenclature. For example, "R", is universal set comprising of all real numbers. The rational numbers, integers and natural numbers are its subset. In other consideration, we can call integers as universal set. In that case, sets such as {1,2,3}, prime numbers, even numbers, odd numbers are subset of the universal set of integers.

The universal set is pictorially represented by a region enclosed within a rectangle on Venn diagram. For illustration, consider the universal set of English alphabets and universal set of first 10 natural numbers as shown in the top row of the figure

Universal set

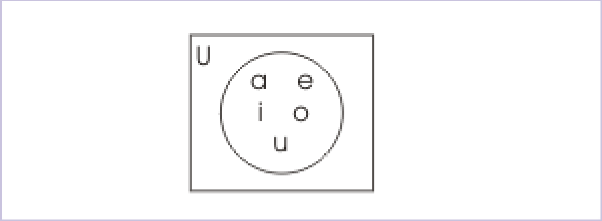


The universal set is represented by a region enclosed within a rectangle.

Many times, however, we may not be required to list elements of a universal set. In such case, we represent the universal set simply by a rectangle and the symbol for universal set, “U”, in the corner. This is particularly helpful, where number of elements in universal set are very large.

The subsets of the universal set are represented by closed curves – usually circles. The subset of vowels (V) is shown here within the circle with the listing of elements. Note that we have not listed all the alphabets for universal set and used the symbol “U” in the corner only.

Subset



The subset of the universal set is represented by a closed curve – usually circle.

Union of sets

Union works on two operands, each of which is a set. The operation is denoted by symbol " ∪ ". Now, the question is : what do we expect when two sets are combined? Clearly, we need to enlist all the elements of two sets in the resulting set.

Union of two sets

> The union of sets “A” and “B” is a third set, which consists all the elements of two sets.

In symbol,

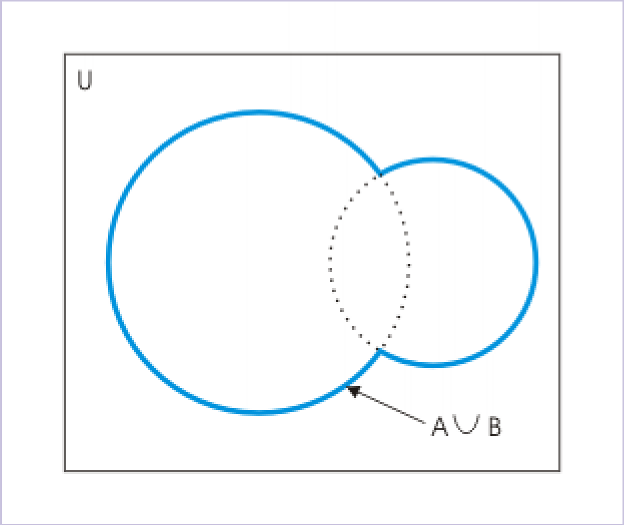
$$A \cup B = \{x : x \in A \quad or \quad x \in B\}$$

The word “or” in the set builder form defining union is important. It means that the element “x” belongs to either “A” or “B”. The element may belong to both sets (common to two sets), but not necessarily. We can, therefore, infer that union set consists of :

1. elements exclusive to A
2. elements exclusive to B
3. elements common to A and B

As a set includes only distinct elements, the common elements are represented only once in the union set. Thus, union set consists of elements of both sets without repeating an element. Now, the set is represented on Venn diagram as shown here.

Union of two sets



The representation of union on Venn diagram

For illustration of working with union, let us consider two sets of positive integers as given here,

$$A = \{1,2,3,4,5,6\}$$

$$B = \{4,5,6,7,8\}$$

The union of two sets is :

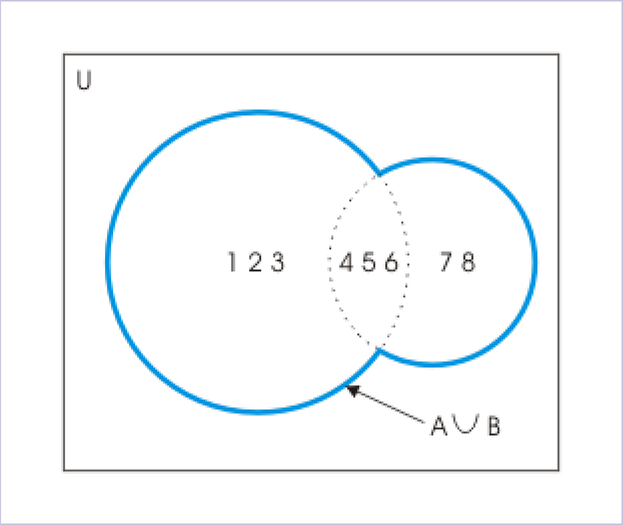
$$\Rightarrow A \cup B = \{1,2,3,4,5,6,4,5,6,7,8\}$$

But repetition of elements in a set does not change it. Hence, we need not repeat elements in the resulting union.

$$\Rightarrow A \cup B = \{1,2,3,4,5,6,7,8\}$$

Here, universal set is natural numbers. The representation of union of joint sets is shown in the figure. We can observe that very construction of union on Venn diagram ensures that elements are not repeated.

Union of two sets



The representation of union on Venn's diagram

Interpretation of union set

Let us examine the defining set of union :

$$A \cup B = \{x : x \in A \quad or \quad x \in B\}$$

We consider an arbitrary element, say “x”, of the union set. Then, we interpret the conditional meaning as :

$$If \quad x \in A \cup B, \quad then \quad x \in A \quad or \quad x \in B.$$

Can we emphasize this conditional meaning in reverse order :

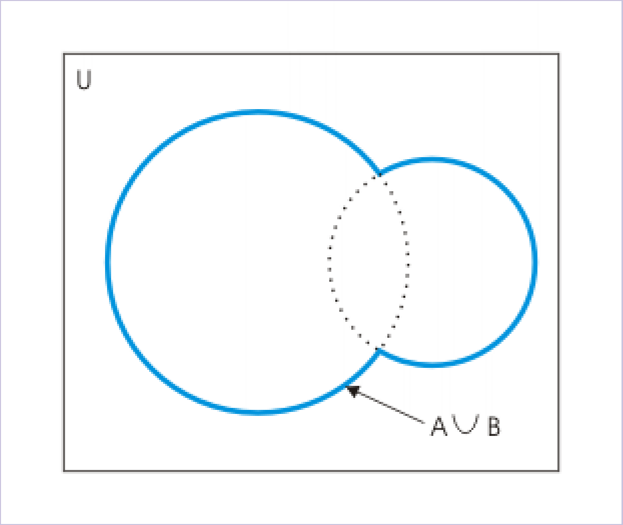
$$If \quad x \in A \quad or \quad x \in B, \quad then \quad x \in A \cup B.$$

Yes, we can agree with the second conditional meaning as well. We, therefore, conclude that the statements work in both ways. We write two statements together as :

$$x \in A \cup B \Leftrightarrow x \in A \quad or \quad x \in B$$

We can reach yet another conclusion by observing representation of union set on Venn diagram. Now, if an arbitrary element “x” does not belong to union set, then it is clear that it does not belong to the region represented by the union set on the Venn’s diagram. Hence,

Union of two sets



The representation of union on Venn's diagram

$$If \quad x \notin \quad \text{(does not belong)} \quad A \cup B \Rightarrow x \notin A \quad \text{and} \quad x \notin B.$$

The important thing to note here is the word “and” in place of “or” used before. Think about it. Here two conditions follow simultaneously. If an element does not belong to an union set, then it will not belong to either of individual sets simultaneously. Now, the next thing to consider is whether this conditional statement will be true other way round as well?

$$If \quad x \notin A \quad and \quad x \notin B \Rightarrow x \notin A \cup B.$$

Yes, we can agree to this statement. We, therefore, conclude that the statements work in both ways. We write two statements together with the help of two ways arrow sign as :

$$x \notin A \cup B \Leftrightarrow x \notin A \quad \text{and} \quad x \notin B.$$

Union of disjoint sets

Consider students in class X and class XI. Let us denote the respective sets as "T" for tenth and "E" for eleventh class. Clearly, union i.e. combination of two sets should include elements from each of the sets. Hence,

$$T \cup E = \text{students in class X and XI}$$

This is a straight forward union of two sets. The resulting set comprises of all elements present in both the sets. Since it is not possible that students studying in class X are also students of XI, we are sure that the numbers of elements in the union is sum of numbers of students in each class. As there is no commonality between two sets, it is a union of two “disjoint” sets. We conclude here that union of two disjoint sets has no common elements.

Union with subset

The set “B” consists of all elements of its subset “A”. In other words, the elements of a subset “A” also belongs to the set “B”. The operation of union is combining elements of two sets. The union with a subset, therefore, consists of elements from both “A” and “B”. However, all elements of “A” are also the elements of “B”. Therefore, we find that union set is same as the superset “B”. Symbolically,

$$If \quad A \subset B, \quad \text{then} \quad B \cup A = B.$$

We can check this deduction with the help of an example. Let us consider two sets as :

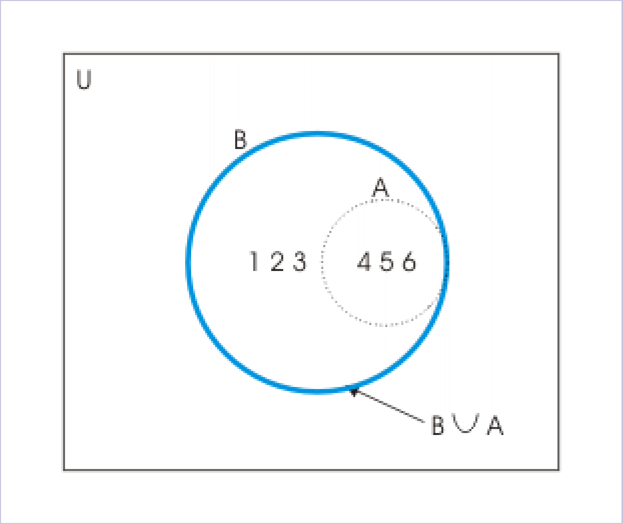
$$A = \{4,5,6\}$$

$$B = \{1,2,3,4,5,6\}$$

Here, we see that A ⊂ B. Now,

$$B \cup A = \{1,2,3,4,5,6,4,5,6\} = \{1,2,3,4,5,6\} = B$$

Union with subset



The representation of union with subset on Venn diagram

Multiple unions

If $A_1, A_2, A_3, \ldots\ldots\ldots, A_n$ is a finite family of sets, then their unions, one after another, is denoted as :

$$A_1 \cup A_2 \cup A_3 \cup \ldots\ldots\ldots \cup A_n$$

Important results

In this section we shall discuss some of the important characteristics/ deductions for the union operation.

Idempotent law

The literal meaning of the word “idempotent” is “unchanged when multiplied by itself”. Following the clue, the union of a set with itself is the set itself. This is an equivalent statement conveying the meaning of “idempotent” in the context of union. Symbolically,

$$A \cup A = A$$

The union set consists of distinct elements and common elements taken once. Between two sets here, all elements are common. The union set consists of all elements of either set.

Identity law

The algebraic operators like addition and multiplication have defined identities, which does not change the other operand of the operator. For example, if we add “0” to a number, it remains same. Hence, “0” is additive identity. Similarly, “1” is multiplicative identity.

In the case of union, we find that union of a set with empty set does not change the set. Hence, empty set is union identity.

$$A \cup \varphi = A$$

As there is no element in empty set, union has same elements as that in “A”.

Law of U

All sets are subsets of universal set for a given context. We have seen that union with subset results in the set itself. Clearly, union of universal set with its subset will result in the universal set itself.

$$U \cup A = U$$

Commutative law

In order to assess whether commutative property holds or not, we consider the example, used earlier. Let the sets be :

$$A = \{1,2,3,4,5,6\}$$

$$B = \{4,5,6,7,8\}$$

Then,

$$A \cup B = \{1,2,3,4,5,6,4,5,6,7,8\} = \{1,2,3,4,5,6,7,8\}$$

$$B \cup A = \{4,5,6,7,8,1,2,3,4,5,6\} = \{1,2,3,4,5,6,7,8\}$$

Thus, we see that order of operands with respect to the union operator is not differentiating. We can also appreciate this law on Venn diagram, which does not change by changing positions of sets across union operator.

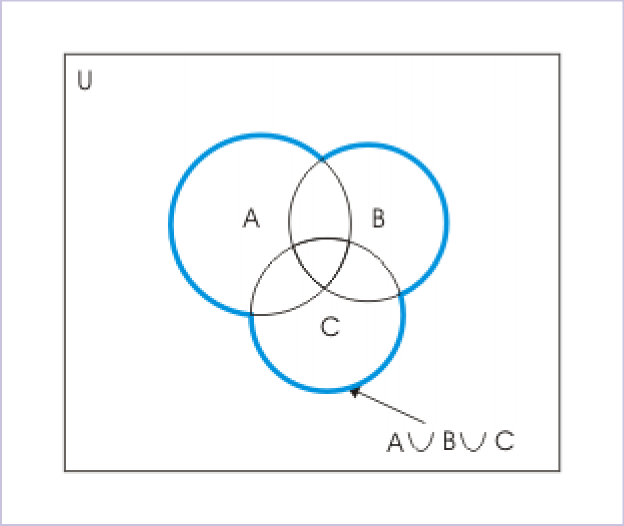
Associative law

The associative property also holds with respect to union operator. We know that associative property is about changing the place of parentheses as here :

$$(A \cup B) \cup C = A \cup (B \cup C)$$

The parentheses simply change the precedence of operation. On Venn diagram, union involving three sets appears same, irrespective of whether we apply union operation in a particular sequence.

Union of three sets



Associative law

Intersection of sets

We have pointed out that a set representing a real situation is not an isolated collection. Sets, in general, overlaps with each other. It is primarily because a set is defined on few characteristics, whereas elements generally can possess many characteristics. Unlike union, which includes all elements from two sets, the intersection between two sets includes only common elements.

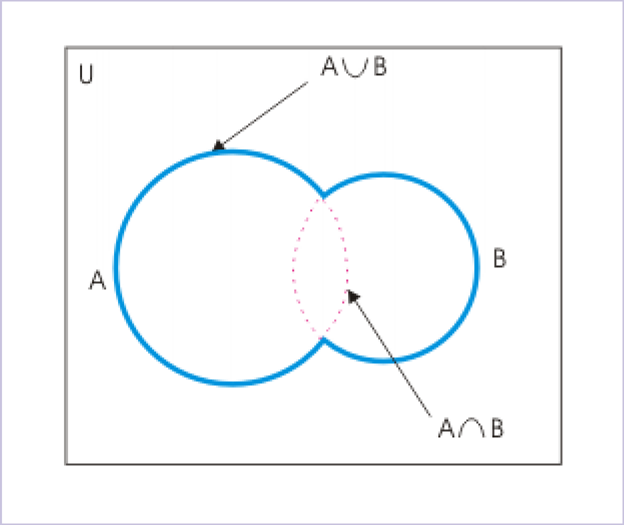
Intersection of two sets

> The intersection of sets “A” and “B” is the set of all elements common to both “A” and “B”.

The use of word “and” between two sets in defining an intersection is quite significant. Compare it with the definition of union. We used the word “or” between two sets. Pondering on these two words, while deciding membership of union or intersection, is helpful in application situation.

The intersection operation is denoted by the symbol, " ∩ ". We can write intersection in set builder form as :

Intersection of two sets



The intersection set consists of elements common to two sets.

$$A \cap B = \{x : \quad x \in A \quad and \quad x \in B\}$$

Again note use of the word “and” in set builder qualification. We can read this as “x” is an element, which belongs to set “A” and set “B”. Hence, it means that “x” belongs to both “A” and “B”.

In order to understand the operation, let us consider the earlier example again,

$$A = \{1,2,3,4,5,6\}$$

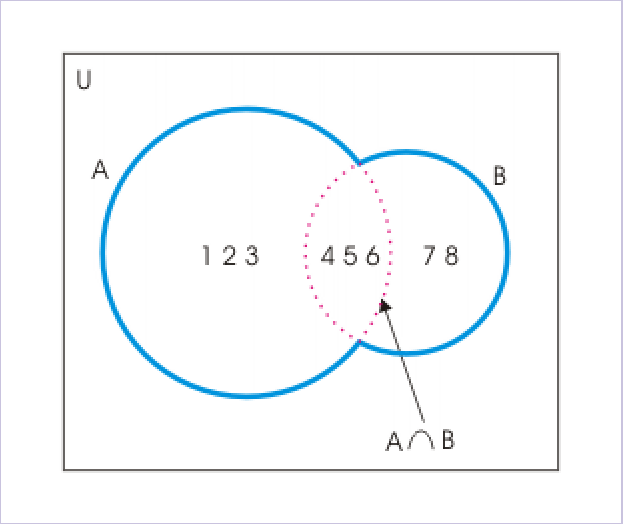
$$B = \{4,5,6,7,8\}$$

Then,

$$A \cap B = \{4,5,6\}$$

On Venn diagram, an intersection is the region intersected by circles, which represent two sets.

Intersection of two sets



The intersection set consists of elements common to two sets.

Interpretation of Intersection set

Let us examine the defining set of intersection :

$$A \cap B = \{x : \quad x \in A \quad and \quad x \in B\}$$

We consider an arbitrary element, say “x”, of the intersection set. Then, we interpret the conditional meaning as :

$$If \quad x \in A \cap B \Rightarrow x \in A \quad and \quad x \in B.$$

The conditional statement is true in opposite direction as well. Hence,

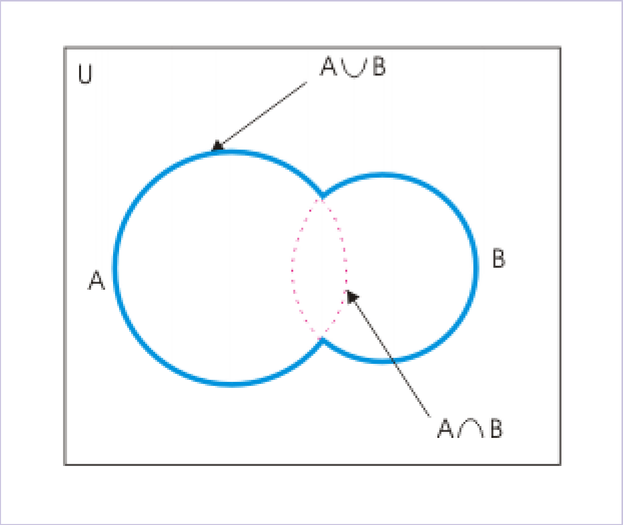
$$If x \in A \quad and \quad x \in B \Rightarrow x \in A \cap B.$$

We summarize two statements with two ways arrow as :

$$x \in A \cap B \Leftrightarrow x \in A \quad and \quad x \in B$$

In addition to two ways relation, there is an interesting aspect of intersection. Intersection is subset of either of two sets. From Venn diagram, it is clear that :

Intersection of two sets



The intersection set consists of elements common to two sets.

$$(A \cap B) \subset A$$

and

$$(A \cap B) \subset B$$

Intersection with a subset

Since all elements of a subset is present in the set, it emerges that intersection with subset is subset. Hence, if “A” is subset of set “B”, then :

$$B \cap A = A$$

Intersection of disjoint sets

If no element is common to two sets “A” and “B” , then the resulting intersection is an empty set :

$$A \cap B = \varphi$$

In that case, two sets “A” and “B” are “disjoint” sets.

Multiple intersections

If $A_1, A_2, A_3, \ldots\ldots\ldots, A_n$ is a finite family of sets, then their intersections one after another is denoted as :

$$A_1 \cap A_2 \cap A_3 \cap \ldots\ldots\ldots \cap An$$

Important results

In this section we shall discuss some of the important characteristics/ deductions for the intersection operation.

Idempotent law

The intersection of a set with itself is the set itself.

$$A \cap A = A$$

This is because intersection is a set of common elements. Here, all elements of a set is common with itself. The resulting intersection, therefore, is set itself.

Identity law

The intersection with universal set yields the set itself. Hence, universal set functions as the identity of the intersection operator.

$$A \cap U = A$$

It is easy to interpret this law. Only the elements in "A" are common to universal set. Hence, intersection, being the set of common elements, is set "A".

Law of empty set

Since empty set is element of all other sets, it emerges that intersection of an empty set with any set is an empty set (empty set is only common element between two sets).

$$\varphi \cap A = \varphi$$

Commutative law

The order of sets around intersection operator does not change the intersection. Hence, commutative property holds in the case of intersection operation.

$$A \cap B = B \cap A$$

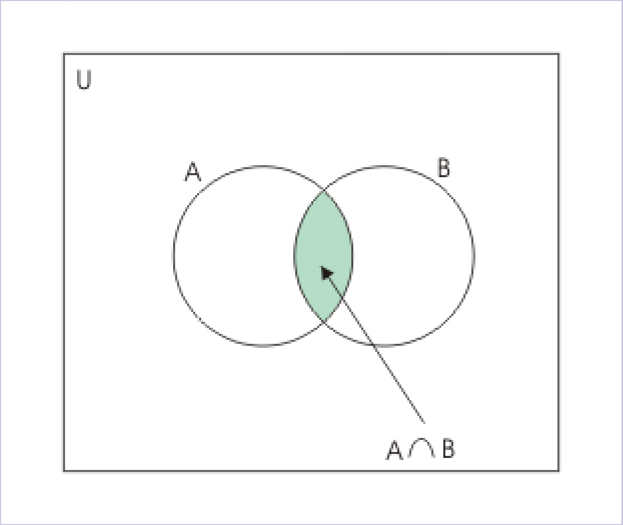
Associative law

The associative property holds with respect to intersection operator.

$$(A \cap B) \cap C = A \cap (B \cap C)$$

The intersection of sets “A” and “B” on Venn’s diagram is :

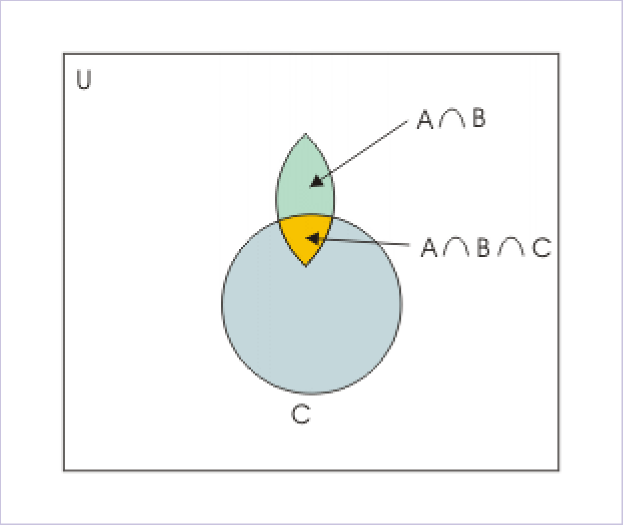
Intersection of two sets



The intersection is a set of common elements and shown as colored region.

In turn, the intersection of set “A ∩ B” and set “C” is the small region in the center :

Intersection inloving three sets



Intersection of a set with "the intersection set of two sets"

It is easy to visualize that the ultimate intersection is independent of the sequence of operation.

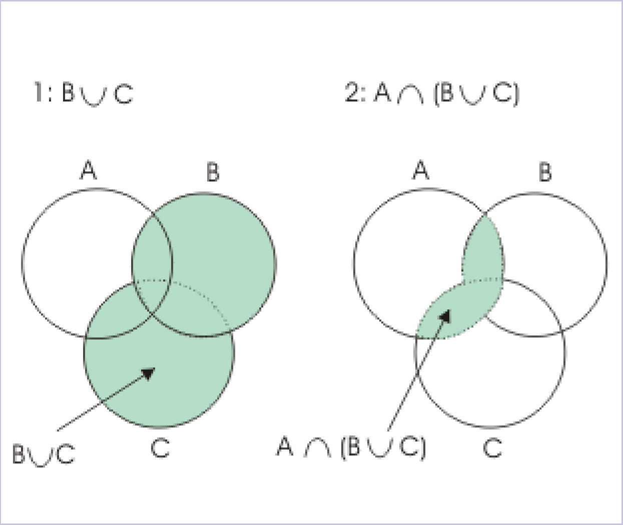
Distributive law

The intersection operator($\cap$) is distributed over union operator ($\cup$) :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

We can check out this relation with the help of Venn diagram. For convenience, we have not shown the universal set. In the first diagram on the left, the colored region shows the union of sets “B” and “C” ie. $B \cup C$. The colored region in the second diagram on the right shows the intersection of set “A” with the union obtained in the first diagram i.e. $B \cup C$.

Distributive law



Distribution of intersection operator over union operator

We can now interpret the colored region in the second diagram from the point of view of expression on the right hand side of the equation :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

The colored region is indeed the union of two intersections : " $A \cup B$" and " $A \cup C$" . Thus, we conclude that distributive property holds for "intersection operator over union operator".

In the same manner, we can prove distribution of “union operator over intersection operator” :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Analytical proof

Distributive properties are important and used for practical application. In this section, we shall prove the same in analytical manner. For this, let us consider an arbitrary element “x”, which belongs to set " $A \cap (B \cup C)$" :

$$x \in A \cap (B \cup C)$$

Then, by definition of intersection :

$$\Rightarrow x \in A \quad and \quad x \in (B \cup C)$$

$$\Rightarrow x \in A \quad and \quad (x \in B \quad or \quad x \in C)$$

$$\Rightarrow (x \in A \quad and \quad x \in B) \quad or \quad (x \in A \quad and \quad x \in C)$$

$$\Rightarrow (x \in A \cap B) \quad or \quad (x \in A \cap C)$$

$$\Rightarrow x \in (A \cap B) \quad or \quad (A \cap C)$$

$$\Rightarrow x \in (A \cap B) \cup (A \cap C)$$

But, we had started with " $A \cap (B \cup C)$ " and used its definition to show that “x” belongs to another set. It means that the other set consists of the elements of the first set – at the least. Thus,

$$\Rightarrow A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$$

Similarly, we can start with " $(A \cap B) \cup (A \cap C)$ " and reach the conclusion that :

$$\Rightarrow (A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$$

If sets are subsets of each other, then they are equal. Hence,

$$\Rightarrow A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Proceeding in the same manner, we can also prove other distributive property of “union operator over intersection operator” :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Difference of sets

We can extend the concept of subtraction, used in algebra, to the sets. If a set "B" is subtracted from set "A", the resulting difference set consists of elements, which are exclusive to set "A". We represent the symbol of difference of sets as "A-B" and pronounce the same as "A minus B".

Difference of sets

> The difference of sets "A-B" is the set of all elements of "A", which do not belong to "B".

In the set builder form, the difference set is :

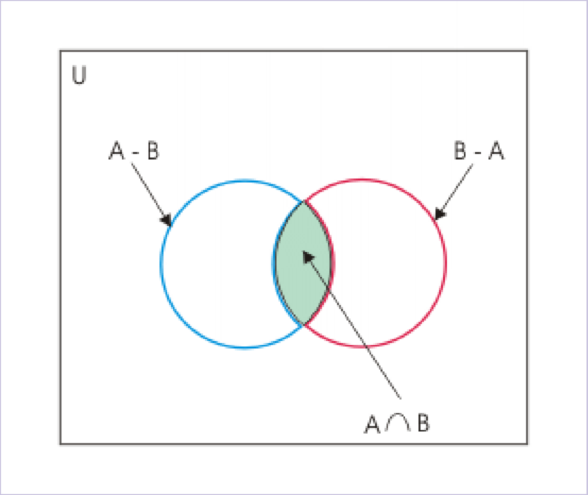
$$A - B = \{x : \quad x \in A \quad and \quad x \notin B\}$$

and

$$B - A = \{x : \quad x \in B \quad and \quad x \notin A\}$$

On Venn's diagram, the difference "A-B" is the region of "A", which excludes the common region with set "B".

Difference of two sets



The difference of two sets is a disjoint set.

Interpretation of difference set

Let us examine the defining set of intersection :

$$A - B = \{x : \quad x \in A \quad and \quad x \notin B\}$$

We consider an arbitrary element, say “x”, of the difference set. Then, we interpret the conditional meaning as :

$$If \quad x \in A - B \Rightarrow x \in A \quad and \quad x \notin B.$$

The conditional statement is true in opposite direction as well. Hence,

$$If \quad x \in A \quad and \quad x \notin B \Rightarrow x \in A - B.$$

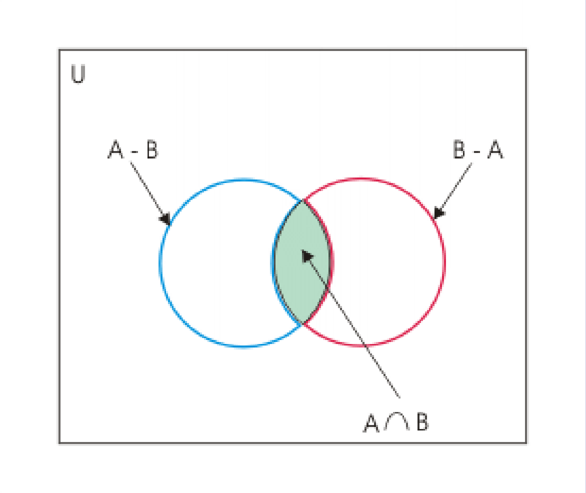
We can summarize two statements with two ways arrow as :

$$If \quad x \in A - B \Leftrightarrow x \in A \quad and \quad x \notin B.$$

Composition of a set

From Venn’s diagram, we observe that if we derive union of $(A \cap B)$ to either of the difference sets, then we get the complete individual set.

Difference of two sets



The difference of two sets is a disjoint set.

$$A = (A - B) \cup (A \cap B)$$

and

$$B = (B - A) \cup (A \cap B)$$

Difference of sets is not commutative

The positions of sets about minus operator affect the result. It is clear from the figure above, where “A-B” and “B-A” represent different regions on Venn’s diagram. As such, the difference of sets is not commutative. Let us consider the example used earlier, where :

$$A = \{1,2,3,4,5,6\}$$

$$A = \{4,5,6,7,8\}$$

Then,

$$\Rightarrow A - B = \{1,2,3\}$$

and

$$\Rightarrow B - A = \{7,8\}$$

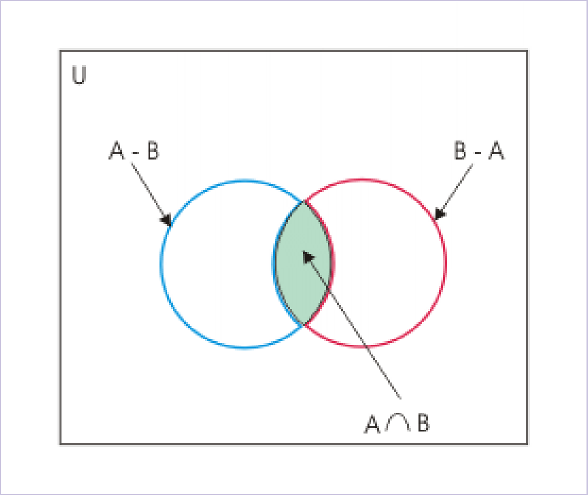
Clearly,

$$\Rightarrow A - B \neq B - A$$

Symmetric difference

From the Venn’s diagram, we can see that union of two sets is equal to three distinct regions. Alternatively, we can say that the region represented by the union of two sets is equal to the sum of the regions representing three “disjoint” sets (i) difference set A-B (ii) intersection set " $A \cap B$" and (iii) difference set B-A.

Difference of two sets



The difference of two sets is a disjoint set.

We use the term “symmetric set” for combining two differences as marked on Venn’s diagram. It is denoted as “ $A \Delta B$”.

$$A \Delta B = (A - B) \cup (B - A)$$

Complement of a set

The complement is a special case of the difference operation. The set in question is subtracted from universal set, “U”. Thus, one of the sets in difference operation is fixed. We define complement of a set as its difference with universal set, "U". The complement of a set is denoted by the same symbol as that of set, but with an apostrophe. Hence, complement of set A is set A’.

Complement of a set

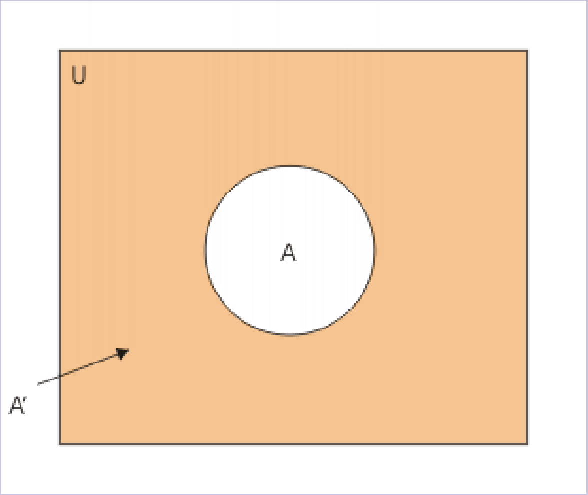
> The complement of a set “A” consists of elements, which are elements of “U”, but not the elements of “A”.

We write the complement set in terms of set builder form as :

$$A' = \{x : \quad x \in U \quad and \quad x \notin A\}$$

Note that elements of A’ does not belong to set “A”. On Venn’s diagram, the complement of “A” is the remaining region of the universal set.

Complement of a set



The complement of a set is the remaining region of the universal set.

Interpretation of complement

Proceeding as before we can read the conditional statement for the complement with the help of two ways arrow as :

$$x \in A\prime \Leftrightarrow x \in U \quad and \quad x \notin A$$

In terms of minus or difference operation,

$$A\prime = U - A$$

It is clear from the representation on Venn’s diagram that the universal set comprises of two distinct sets – set A and complement set A’.

$$\Rightarrow U = A \cup A\prime$$

Compliment of universal set

The complement of universal set is empty set. It is so because difference of union set with itself is the empty set (see Venn's diagram).

$$U\prime = \{x : \quad x \in U \quad and \quad x \notin U\} = \varphi$$

Complement of empty set

The complement of the empty set is universal set. It is so because difference of union set with the empty set is universal set (see Venn's diagram).

$$\varphi\prime = \{x : \quad x \in U \quad and \quad x \notin \varphi\} = U$$

Complement of complement set is set itself

The complement of complement set is set itself. The complement set is defined as :

$$A\prime = U - A$$

Now, complement of complement set is :

$$\Rightarrow (A\prime)\prime = (U - A)\prime$$

Let us consider the example, where :

$$U = \{1,2,3,4,5,6,7,8\}$$

$$A = \{1,2,3,4,5,6\}$$

Then,

$$\Rightarrow A\prime = \{1,2,3,4,5,6,7,8\} - \{1,2,3,4,5,6\} = \{7,8\}$$

Again taking complement, we have :

$$(A\prime)\prime = \{1,2,3,4,5,6,7,8\} - \{7,8\} = \{1,2,3,4,5,6\} = A$$

Union with complement set

The union of a set with its complement is universal set :

$$A \cup A\prime = \{x : \quad x \in U \quad and \quad x \in A\} \cup \{x : \quad x \in U \quad and \quad x \in A\} = U$$

From Venn's diagram also, we see that universal set consists of set A and component A'.

$$U = A \cup A\prime$$

The two sets on the right side of the equation are disjoint sets. Hence,

$$A \cup A\prime = U$$

Intersection with complement set

There is nothing common between set A and its component A'. Thus, intersection of a set with its complement yields the empty set,

$$A \cap A\prime = \varphi$$

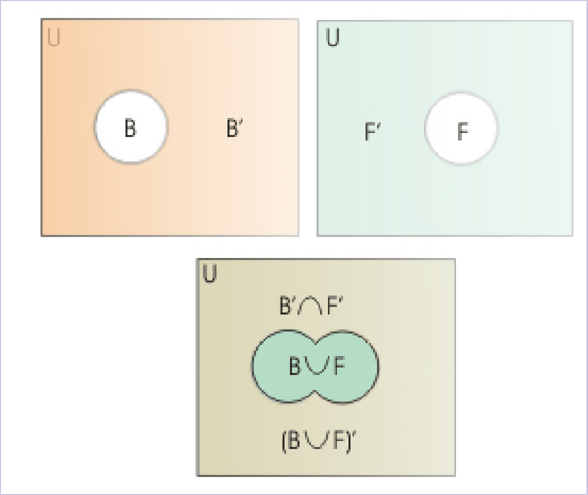
De-morgan's laws

In the real world situation, we want to negate a condition of incidence. For example, consider a class in the school. Some students play either basketball or football or both, but there are students, who play neither basketball nor football. We have to identify later category of students as a set.

Let the set of students playing basketball be "B" and that playing football be "F". Then, students who do not play basketball is complement set B' and students who do not play football is complement set F'. We have shown these complement sets separately for visualization. Actually, these complement sets are drawn to the same universal set, "U".

Two complement sets are but overlapping sets. There are students in the set B' who play football and there are students in the set F', who play basketball. In order to remove those students playing other game, we intersect two complements. The members of the intersection of two complements, therefore, represent students who play neither basketball nor football. This intersection is shown as third bottom Venn’s diagram in the figure.

Intersection



Intersection of two component sets

Looking at the intersection of two complement sets, however, we observe that this is equal to the complement of union “ $B \cup F$ ”. This conclusion can be derived from basic interpretation as well. We know that union “ $B \cup F$ ” represents students, who play either or both games. The component of the union, therefore, represents, who neither play basketball nor football.

This fact, as a matter of fact, is the first De-morgan’s law. Symbolically,

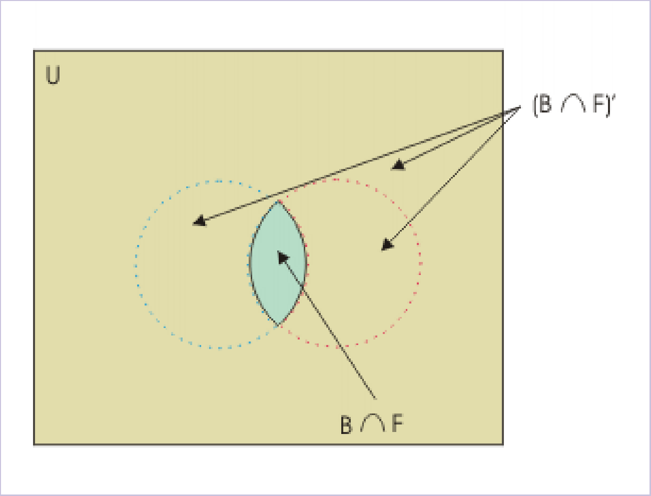
$$B\prime \cap F\prime = (B \cup F)\prime$$

The second De-morgan’s law is :

$$B\prime \cup F\prime = (B \cap F)\prime$$

In the parlance of illustration given earlier, let us interpret right hand side of the second De-morgan's law. The intersection “ $B \cap F$ ” represents students playing both games. Its complement, therefore, represents students who do not play both games, but may play one of them.

Component set



Component of intersection of two sets

Analytical proof

Here, we shall prove first De-morgan's law in this section. The second law can be proved in similar fashion. Let us consider an arbitrary element "x" belonging to set ($A \cup B$)'.

$$x \in (A \cup B)'$$

$$\Rightarrow x \notin (A \cup B)$$

Then, by definition of union,

$$\Rightarrow x \notin \{x : \quad x \in A \quad or \quad x \in B\}$$

Here, "not or" is interpreted same as "and",

$$\Rightarrow x \notin A \quad and \quad x \notin B$$

$$\Rightarrow x \in A' \quad and \quad x \in B'$$

$$\Rightarrow x \in A' \cap B'$$

But, we had started with ($A \cup B$)’ and used its definition to show that “x” belongs to another set. It means that the other set consists of the elements of the first set – at the least. Thus,

$$(A \cup B)' \subset A' \cap B'$$

Similarly, we can start with $A' \cap B'$ and reach the conclusion that :

$$A' \cap B' \subset (A \cup B)'$$

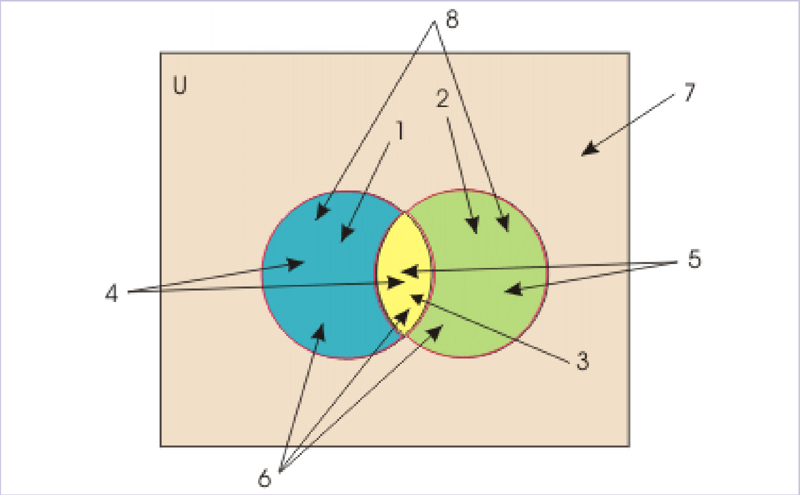
If sets are subsets of each other, then they are equal. Hence,

$$A' \cap B' = (A \cup B)'$$

Example

Problem 1: In the reference of students in a class, the set “B” represents students, who play basketball. The set “F” represents students, who play football. The set “B” and “F” are left and right circles respectively on the Venn's diagram shown below. Identify regions marked 1 to 8 on the Venn’s diagram. Also interpret regions identified by combination U – (6+7).

sets



Interpreting sets

Solution : The meaning of regions market 1 – 8 are as given hereunder :

1 : B-F : It represents the difference of “B” and “F”. It consists of students, who play basketball, but not football.

2 : F-B : It represents the difference of “F” and “B”. It consists of students, who play football, but not basketball.

3 : $B \cap F$: It represents the intersection of two sets. It consists of students, who play both basketball and football.

4 : B: It represents the set “B”. It is union of two disjoint sets “B-F” and “ $B \cup F$ ”. It consists of students, who play basketball.

5 : F: It represents the set “F”. It is union of two disjoint sets “F-B” and “ $B \cap F$ ”. It consists of students, who play football.

6 : B∪F: It represents the union set of set “B” and “F”. Equivalently, it is union of three disjoint sets “B-F”, “ $B \cap F$ ” and “F-B”. It consists of students, who play either of two games or both.

7 : ($B \cup F$)’: It represents the component of union set “ $B \cup F$ ”. It consists of students, who play neither basketball nor football.

8 : $(B - F) \cup (F - B)$: It represents union of two disjoint difference sets “B-F” and “F-B”. It consists of students, who play only one game.

The region, identified by U – (6+7), is complement of “ $B \cap F$ ”. It represents students, who do not play both games, but may play one of them.

Working with two sets

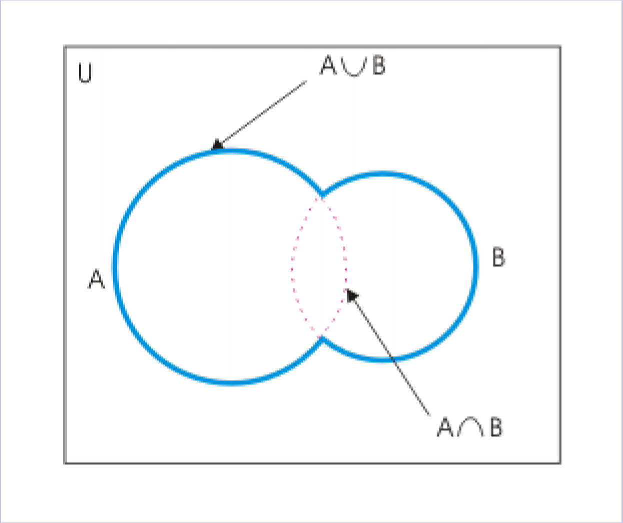
There are finite numbers of elements in finite set. This allows us to analyze numbers of elements in different sets that results from the operations carried on them. In this module, we shall study different operations on sets in the context of practical applications. However, we shall limit ourselves to the interaction, involving two sets. The interaction, involving three sets, will be dealt in a separate module.

We use a specific notation to represent the numbers of elements in a set. For example, the numbers in set "A" is represented as "n(A)", whereas we denote numbers of elements in the union as "n($A \cup B$)".

Elements in the union of two sets

The area, demarcated with solid line, in the Venn's diagram, shows the union of two sets denoted by ($A \cup B$). We want to know the numbers of elements in this union in terms of numbers of elements in individual sets.

Union of two sets



The common elements in the union set is counted only once.

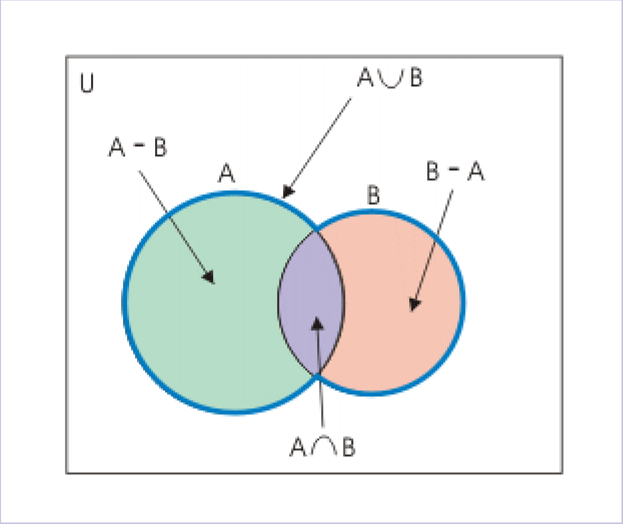
The sum of the numbers in the individual sets is generally greater than the numbers in the union. The reason is that union includes common elements only once. On the other hand, sum of the numbers of individual sets counts common elements once with each set – in total two times. Clearly, it is required that we deduct the numbers of elements, which are common to each set, from the sum of numbers of elements in individual sets. Hence,

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

Here, n($A \cap B$) represents the numbers of elements common to two sets. As reminder only, we note that plus (+) operation is not a valid set operation. We, however, use this algebraic operation here as we are now dealing with the numbers in set - not the set.

Alternatively, we can approach this expansion in yet another way. See the representation of intersection of two sets. The union of two sets can be considered to comprise of three distinct regions. Three regions shown with different colors represent three "disjoint" sets. Clearly,

Union of two sets



The region representing union consists of three distinct or disjointed region.

$$n(A \cup B) = n(A - B) + n(A \cap B) + n(B - A)$$

However, we observe that if we add n($A \cup B$) to either of the two difference sets, then we get the complete individual set.

$$n(A) = n(A - B) + n(A \cap B)$$

$$\Rightarrow n(A - B) = n(A) - n(A \cap B)$$

and

$$n(B) = n(B - A) + n(A \cap B)$$

$$\Rightarrow n(B - A) = n(B) - n(A \cap B)$$

Substituting for the numbers of the difference set in the equation for the numbers in the union set, we have :

$$\Rightarrow n(A \cup B) = n(A) - n(A \cap B) + n(A \cap B) + n(B) - n(A \cap B)$$

$$\Rightarrow n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

Numbers of elements in the union of “disjoint” sets

Since there are no common elements between two disjoint sets, the intersection between disjoint sets is an empty set. Hence,

$$n(A \cup B) = n(A) + n(B)$$

Application

Application of set theory to real situation is keyed to the interpretation of wordings and description. In order to efficiently employ the concepts of set theory to real world situations, we need to interpret description of collection appropriately.

Once collections are interpreted correctly, rest is easy. There are indeed fewer mathematical operations involved here. Most of these relate to determination of numbers of elements in a set. In this section, we shall first recapitulate or reinterpret different collections and then work with few representative situations for analysis (if we are confident then we can skip the recapitulation part).

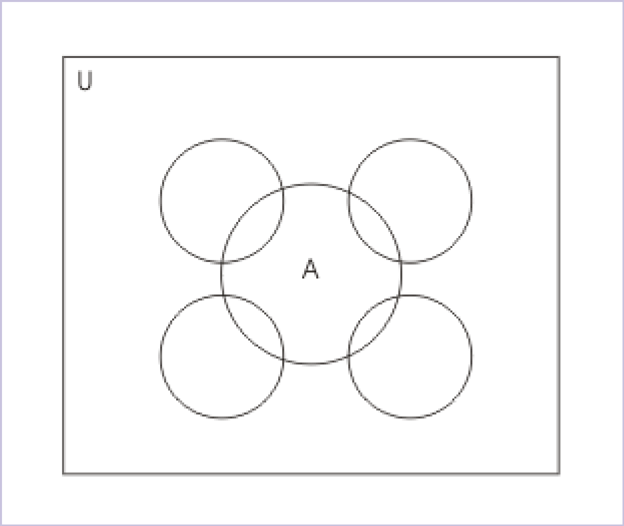
Set

In real situation, we identify a collection with certain characteristic common to elements. For example, a set of students in a class is based on the characteristic that each student is member of that class. This type of interpretation, however, is generally restrictive and leads to misinterpretation. We tend to think that the collection is isolated in itself, which is obviously wrong.

We need to free our mind from thinking set as an isolated entity. Some of the students might be members of another collection like that of basketball team, whereas some others might be members of a particular house, say “Amity house” and so on

In the nutshell, we consider set as a collection, which has multiple intersections with other collections.

A set



A set has multiple intersections
with other collections.

Example

Problem 1: In the house of total 200 students, 140 students play basketball and 80 students play football. Each student of the house plays at least one of these two games. How many students play both basketball and football?

Solution : The individual sets here are students playing basket ball (B) and football (F). Hence,

$$n(B) = 140$$

$$n(F) = 80$$

Clearly, there is no bar that a students playing basketball can not play football. This is also evident from the sum of the numbers in each set. The sum is 140 + 80 = 220, whereas total numbers of students in the house is 200 only. Thus, there are students who play both games. We can interpret the total numbers as the union of two individual sets. Hence, applying expansion for the numbers of a union :

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

The students who play both games constitute the intersection of two individual sets.

Putting values,

$$\Rightarrow n(B \cup F) = 140 + 80 - 200 = 20$$

Universal set and complement

Universal is inclusive of all related sets. If we observe the Venn’s diagram consisting of two individual sets, then we realize that largest closed region within the universal set is the union involving two sets i.e (A∪B). This union, however, is a subset of U. There is remaining area within the universal set, which is called the component of this union.

Now we know that a union represents elements which belong to either set exclusively or belong commonly with other sets. It means that the complement of union represents the region, which can not be defined by the characterizing criteria of the union. This complement of union, therefore, represents situations which is described in terms of “neither or nor” type. Actually, this set is given by De-morgan’s first law.

Example

Problem 2: In a house of total 200 students, 100 students play basketball, 60 students play football and 20 play both games. How many students play neither basketball nor football?

Solution : We have already discussed that “neither nor” condition is same as that of De-morgan’s first law :

$$n(B' \cap F') = n(B \cup F)'$$

Now expanding the right hand term, we have :

$$\Rightarrow n(B' \cap F') = n(B \cap F)' = U - n(B \cup F)$$

Further using formula for the numbers in a union,

$$\Rightarrow n(B' \cap F') = U - n(B) - n(F) + n(B \cap F)$$

Putting values,

$$\Rightarrow n(B' \cap F') = 200 - 100 - 60 + 20 = 60$$

This is the required answer. However, there remains a question : why do we consider total numbers of students as the numbers in universal set, “U”, unlike previous example in which this number corresponds to numbers in the union of individual sets. Remember, earlier question had the phrase “Each student of the house plays at least one of these two games”. This ensured that total numbers represented the union as everyone was playing one of two games. Such restriction is not there in this example. In fact, we saw that there are students who are not playing either of two games at all! Thus, total number represents universal set in this example.

Union

Union of two sets “A” and “B” conveys the meaning of consisting three categories of elements (i) elements exclusively belonging to “A” (ii) elements exclusively belonging to “B” and (iii) (i) elements commonly belonging to “A” and “B”. In totality, we see that union conveys the meaning of “or” – the elements may belong either to a particular set or to both sets.

Example

Problem 3: In a group of students, 40 students study either English or Mathematics. Of these 25 students study Mathematics, 10 students study both Mathematics and English. How many students study English?

Solution : The word “or” in the first sentence indicates that union of students studying Mathematics (M) or English (E) or both is 40. Using formula, we have :

$$n(M \cup E) = n(M) + n(E) - n(M \cap E)$$

$$\Rightarrow n(E) = n(M \cup E) - n(M) + n(M \cap E)$$

Putting values,

$$\Rightarrow n(E) = 40 - 25 + 10 = 25$$

Difference

In the case of intersection of two sets, we have noted that difference represents the exclusive or isolated set, which is not common to other set. From the Venn's diagram, we also observe that a given set is actually composed of two sets (i) difference set and (ii) intersection set.

$$n(A) = n(A - B) + n(A \cap B)$$

and

$$n(B) = n(B - A) + n(A \cap B)$$

Example

Problem 4: In a house of 200 students, 120 students study Mathematics, 60 students study English and 40 students study both Mathematics and English. Find in the house : (i) students who study Mathematics but not English (ii) students who study English, but not Mathematics (iii) students who study either Mathematics or English and (iv) students who neither study Mathematics nor English.

Solution : Let us first characterize collections as given in the question. Two sets are given one for those who study Mathematics (M) and other for those who study English(E). The addition of numbers of individual sets is 120 + 60 = 180, which is less than total numbers of students. Hence, total numbers of 200 corresponds to universal set. Here,

$$U = 200; \quad n(M) = 120; \quad n(E) = 60 \quad and \quad n(M \cap E) = 40.$$

(i) Students studying Mathematics, but not English means that we need to find the numbers in the difference of set i.e M – E.

$$n(M - E) = n(M) - n(M \cap E)$$

$$\Rightarrow n(M - E) = 120 - 40 = 80$$

(ii) Students studying Mathematics, but not English means that we need to find the numbers in the difference of set i.e E – M.

$$n(E - M) = n(E) - n(M \cap E)$$

$$\Rightarrow n(E - M) = 60 - 40 = 20$$

(iii) Students who study either Mathematics or English is equal to the numbers in the union of two sets.

$$n(M \cup E) = n(M) + n(E) - n(M \cap E)$$

Putting values,

$$\Rightarrow n(M \cup E) = 120 + 60 - 40 = 140$$

(iv) Students who study neither Mathematics nor English is equal to the numbers in the compliment of the union of two sets.

$$n(M \cup E)' = U - n(M \cup E) = 200 - 140 = 60$$

Working with three sets

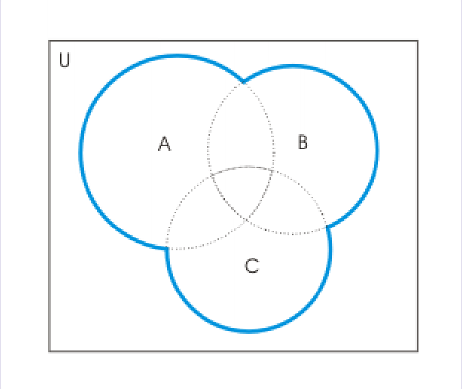
Working with three sets is similar as working with two sets. The underlying characteristics of set operations are same. The union and intersection of sets are carried out with three sets - one after another. Notably, union and intersection operations are commutative. This allows us to extend these set operations to third set in any sequence. Venn's diagram enables us to visualize resulting set. In this module, we shall first formulate expression for the numbers in the union of three sets. Subsequently, we shall apply the formulation to real time analysis - using both graphical (Venn diagram) and analytical methods.

The union, involving three sets, can be considered in terms of union of a set with "union of other two sets". In that sense, union of three sets represent elements which belong to either of three sets. Here, we want to find the expression of numbers of elements in the union set, which is represented as :

$$n(A \cup B \cup C)$$

Before, we work on the expansion of this term, let us first find out what does the term " $A \cup B \cup C$ " represent on Venn's diagram? The figure below shows the representation of this term :

Union of three sets



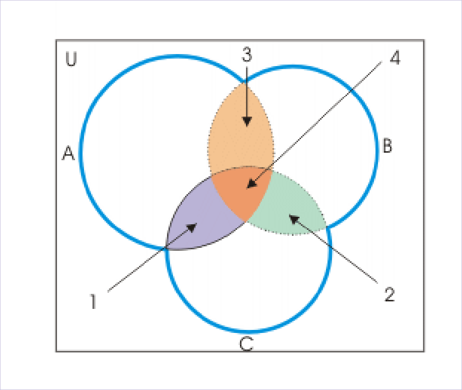
The union set represent elements of three sets combined together.

The set shown in the figure above consists of following class of elements :

1. The elements, which are exclusive of sets "A", "B" and "C" respectively.
2. The elements, which are common to a pair of two sets at a time.
3. The elements, which are common to all three sets.

In the figure below, the common areas between a pair of two sets are marked "1", "2" and "3". The common area among all three sets is marked "4".

Union of three sets



The union consists of disjointed regions.

Union of three sets

As discussed earlier, the sum of numbers of individual sets is greater than the number of elements in " $A \cup B \cup C$ ", unless the sets are disjoint sets. It is imperative that we account for the repetition of common elements. Proceeding as in the case of union of two sets, we deduct the intersections between each pair of sets as :

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C)$$

In this manner, we account for common elements between two sets. However, we have deducted elements "common to all three sets" in this process – three times. On the other hand, the elements "common to all three sets" are present in the numbers of each of the individual sets - in total three times as there are three sets. Ultimately, we find that we have not counted the elements common to all sets at all. It means that we need to account for the elements common to all three sets. In order to add this number, we first need to know – what does this common area (marked 4) represent symbolically?

In the earlier module, we have seen that the area marked "4" is represented by " $A \cap B \cap C$ ". Hence, the correct expansion for the numbers of elements in the union, involving three set, is :

$$\Rightarrow n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$$

Note: This result is an important result as the same is used while studying probability.

Union of three sets (Analytical method)

We can achieve this result analytically as well. Here, we consider "A" as one set and " $(B \cup C)$ " as other set. Then, we apply the relation, which has been obtained for the numbers in the union of two sets as :

$$n(A \cup B \cup C) = n(A) + n(B \cup C) - n[A \cap (B \cup C)]$$

Applying result for the union of two sets for " $n(B \cup C)$ ", we have :

$$n(B \cup C) = n(B) + n(C) - n(B \cap C)$$

Putting in the expression for " $n(A \cup B \cup C)$ ",

$$\Rightarrow n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(B \cap C) - n[A \cap (B \cup C)]$$

At this stage, our task is to evaluate " $n[A \cap (B \cup C)]$ ", Recall that we have worked with the distributive property of "intersection operator over union operator". Following distributive property,

$$\Rightarrow n[A \cap (B \cup C)] = n[(A \cap B) \cup (A \cap C)]$$

We can treat each of the terms in the small bracket on the right hand side of the above equation as a set. Applying relation obtained for the numbers in the union of two sets again, we have :

$$\Rightarrow n[A \cap (B \cup C)] = n(A \cup B) + n(A \cup C) - n[(A \cap B) \cap (A \cap C)]$$

The last term in above equation is :

$$[(A \cap B) \cap (A \cap C)] = (A \cap B \cap C)$$

Hence,

$$\Rightarrow n[A \cap (B \cup C)] = n(A \cup B) + n(A \cup C) - n[(A \cap B \cap C)]$$

Now, putting this expression in the expression of the numbers in the union involving three sets and rearranging terms, we have :

$$\Rightarrow n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$$

In the nutshell, we find that numbers of elements in the union, here, is equal to the sum of numbers in the individual sets, minus elements common to two sets taken at a time, plus elements common to all three sets.

Illustration

In this section, we shall work with an example, which is quite intuitive of the analysis, involving three sets. We shall see that analysis of set operations in terms of Venn's diagram is very direct and simple. As such, we shall first attempt analyze situation with Venn's diagram.

However, we need to emphasize that extension of set concepts to calculus, probability and other branches of mathematics require that we develop analytical skill with respect to set operations. Keeping this aspect in mind, we shall also work the solution, using analytical method.

Problem : In a town, a total 100000 people read newspaper. Out of these, 40 % read newspaper "A", 30 % read newspaper "B", 10 % read newspaper "C". It is found that 5% read both "A" and "B"; 4% read both "A" and "C"; and 3% read both "B" and "C". Also, 2% of the people read all three newspapers. Find numbers (i) who read only "A" (ii) who read only "B" (iii) who read neither of three newspapers.

We define three sets “A”, “B” and “C”, corresponding to people reading newspapers “A”, “B” and “C” respectively. From question, we have :

$$n(A) = 0.4X100000 = 40000$$

$$n(B) = 0.3X100000 = 30000$$

$$n(C) = 0.1X100000 = 10000$$

$$n(A \cap B) = 0.05X100000 = 5000$$

$$n(A \cap C) = 0.04X100000 = 4000$$

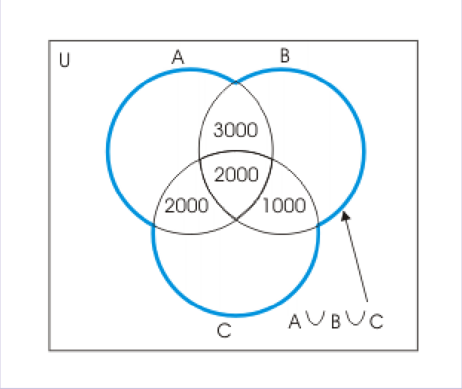
$$n(B \cap C) = 0.03X100000 = 3000$$

$$n(A \cap B \cap C) = 0.2X100000 = 2000$$

Venn's diagram method

We observe that sum of the individual sets is less than total numbers of people , reading newspaper, in the town. Hence, total reading population represents universal set, “U”. The representations of these sets are shown on Venn’s diagram. Note that we have split the elements common to a pair of two sets in two parts (a) elements exclusive to intersection of two sets and (b) elements common to all three sets.

Union of three sets



The regions common to sets.

From the diagram,

(i) The required set is the region of “A” not common to “B” and “C”. This region represents elements, which are exclusive to set “A”. Thus, numbers of people reading only “A” is :

$$n_1 = 40000 - (3000 + 2000 + 2000) = 33000$$

(ii) The required set is the region of “B” not common to “A” and “C”. This region represents elements, which are exclusive to set “B”. Thus, numbers of people reading only “B” is :

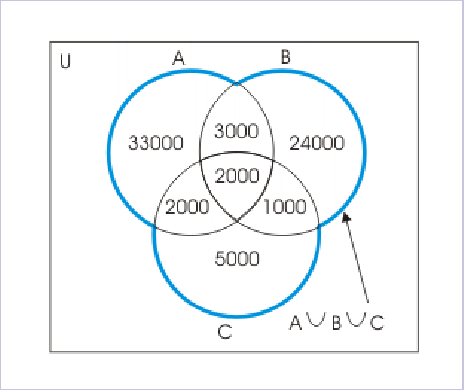
$$n_2 = 30000 - (3000 + 1000 + 2000) = 24000$$

(iii) The required set is the remaining region of universal set “U” i.e. complement of the union of three sets. Now, proceeding as before, people who read newspaper “C” only is (see Venn’s diagram above):

$$n_3 = 100000 - (2000 + 1000 + 2000) = 5000$$

Hence, the required number is :

Union of three sets



The regions representing people, who read neither of three newspapers.

$$n_4 = 100000 - (33000 + 24000 + 5000 + 3000 + 2000 + 1000 + 2000)$$

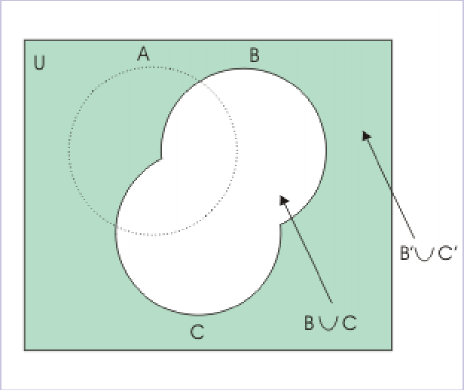
$$\Rightarrow n_4 = 30000$$

Analytical method

(i) Here we are required to find the numbers of people reading only “A”. It is clear that this set is part of the people, who do not read newspapers “B” and “C”. As discussed in the case of two sets, the numbers of people who read neither “B” and “C” is given by De-morgan’s first equation,

$$B' \cap C' = (B \cup C)'$$

Intersection of two complement sets



The region representing people, who read neither of two newspapers.

For clarity, we have shown this region in the Venn’s diagram. We should realize that this intersection of two sets also includes people who read newspaper “A”. However, we are required to know numbers of people, who read newspaper “A” only. The exclusion people reading other newspaper as well are not part of our required set.

The remaining set (refer Venn’s diagram) represent area consisting of people who exclusively read newspaper “A” or who do not read any of three newspapers. Now, we need the intersection of set A with the remaining region to obtain the numbers, who read only newspaper “A”. Hence, required number is :

$$n_1 = n(A \cap B' \cap C')$$

Using De-morgan's equation,

$$n_1 = n(A \cap B' \cap C') = n[A \cap (B \cup C)']$$

$$\Rightarrow n_1 = n[A \cap [U - (B \cup C)]] = n[(A \cap U) - A \cap (B \cup C)]]$$

$$\Rightarrow n_1 = n[A - A \cap (B \cup C)] = n(A) - n[A \cap (B \cup C)]$$

Using distributive property of intersection over union,

$$\Rightarrow n_1 = n(A) - n[(A \cap B) \cup (A \cap C)]$$

Using formula of expansion of union of two sets,

$$\Rightarrow n_1 = n(A) - [n(A \cap B) + n(A \cap C) - n[(A \cap B) \cap (A \cap C)]]$$

$$\Rightarrow n_1 = n(A) - [n(A \cap B) + n(A \cap C) - n(A \cap B \cap C)]$$

We see that values of each term on the right hand side are given. Putting these values,

$$\Rightarrow n_1 = 40000 - [5000 + 4000 - 2000] = 33000$$

(ii) Here we are required to find the numbers of people reading only “B”. Proceeding as before,

$$n_2 = n(B) - [n(B \cap A) + n(B \cap C) - n(B \cap A \cap C)]$$

$$\Rightarrow n_2 = n(B) - [n(A \cap B) + n(B \cap C) - n(A \cap B \cap C)]$$

Putting values,

$$\Rightarrow n_2 = 30000 - [5000 + 3000 - 2000] = 24000$$

(iii) In order to find the numbers of people, who read neither of three newspapers, we first find union of three sets. The union represents people who read either of these newspapers – one, two or all three. Clearly, people, who do not read either of these papers constitute a complement of this union.

$$n_3 = (A \cup B \cup C)' = U - (A \cup B \cup C)$$

Using expansion for the numbers in the union of three sets,

$$\Rightarrow n_3 = U - [n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)]$$

Putting values, we have :

$$\Rightarrow n_3 = 100000 - [40000 + 30000 + 10000 - 5000 - 4000 - 3000 + 2000] = 30000$$

Cartesian product

We have seen that set operations convey the notion of arithmetic operations. One such similar operation is product of two sets called “Cartesian product”. Since sets are collection – not a single quantity, the product operation here involves combining or pairing each of the elements of one set with that of another set.

We use symbol “X” to denote product operation. The Cartesian product of two sets “A” and “B” is symbolically represented as :

$$A \times B$$

It is important to understand that we do not multiply elements as we do in arithmetic – instead we pair elements together. This is the meaning of “product” for the sets. We denote one such pair within a pair of small brackets like :

$$(a, b)$$

where $a \in A$ and $b \in B$.

Note that elements from two sets are separated by comma.

Ordered pair

The order of pairing is important. The pair (a,b) and (b,a) are different. This ordering is required as there are real time situations, where order makes a difference. Consider for example, we are required to find the integers which can be formed from two integer subsets like {1,2,3} and {3,4,5}. Clearly, “13” and “31” represent different integers. We need to distinguish them. All pairs formed from two sets should be distinct.

Keeping this restriction in mind, let us work out an example to find ordered pairs formed from elements of two sets.

$$A = \text{set of first letter of the names of cities} = \{N, D, H\}$$

$$B = \text{set of numbers denoting flight numbers} = \{001{,}002{,}003\}$$

All possible ordered pairs formed from two sets are :

$$(N, 001), (N, 002), (N, 003), (D, 001), (D, 002), (D, 003), (H, 001), (H, 002), (H, 003)$$

There are all together 9 ordered pairs. From this example, we can deduce a method for writing ordered pairs from two sets. We begin with the first elements of two sets. Progressively, we change the elements from the second set till it is exhausted, while keeping the elements from the first set unchanged. Then, we switch to “next” element from first set and start with “first” from the second set. Again, we change the elements from the second set progressively till it is exhausted, while keeping the elements from the first set unchanged. We continue in this manner till all elements from the first set is also exhausted.

From this discussion, it is also evident that two ordered pairs are equal if and only if the corresponding first and second elements are equal.

Cartesian product

The Cartesian product of two sets is defined in terms of ordered pairs.

Cartesian product

> The Cartesian product of two non-empty sets “A” and “B” is the set of all ordered pairs of the elements from two sets.

We should emphasize the use of word “non-empty”, The Cartesian product of a non-empty set with an empty set is equal to empty set.

$$A \times \varphi = \varphi$$

On the other hand, if one of the sets is infinite, then resulting Cartesian product is also infinite.

We express the Cartesian product set in set building form as :

$$A \times B = \{(x, y) : \quad x \in A, y \in B\}$$

Here, use of "," in the set builder form is equivalent to "and". Therefore, we can write Cartesian product of two sets also as :

$$A \times B = \{(x, y) : \quad x \in A \text{ and } y \in B\}$$

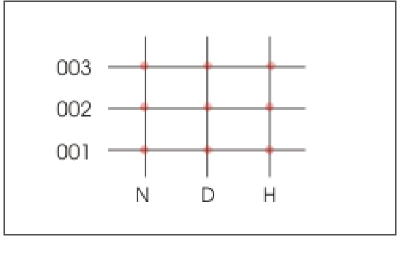
Further, we can emphasize two ways validity of the conditional statements as in the case of other set operators :

$$\text{If}(x, y) \in A \times B \Leftrightarrow x \in A \text{ and } y \in B$$

Graphical representation

The ordered pairs can be represented in the form of tabular cells or points of intersection of perpendicular lines. The elements of one set are represented as rows, whereas elements of other set are represented as columns. Look at the representation of ordered pairs by points in the figure for the example given earlier.

Cartesian product



The elements of one set
are represented as rows,
whereas elements of
other set are represented
as columns.

Note that there are a total of 9 intersection points, corresponding to 9 ordered pairs.

Examples

Problem 1 : If $\left(x^2-1, y+2\right)=(0,2)$, find “x” and “y”.

Solution : Two ordered pairs are equal. It means that corresponding elements of the ordered pairs are equal. Hence,

$$\Rightarrow x^2-1=0$$

$$\Rightarrow x=1 \quad or \quad -1$$

and

$$\Rightarrow y+2=2$$

$$\Rightarrow y=0$$

Problem 2 : If A = {5,6,7,2}, B={3,5,6,1} and C = {4,1,8}, then find $(A \cap B) \times (B \cap C)$.

Solution : In order to evaluate the given expression, we first find out the intersections given in the brackets.

$$\Rightarrow A \cap B=\{5,6\}$$

$$\Rightarrow B \cap C=\{1\}$$

Thus,

$$(A \cap B) \times (B \cap C)=\{(6,1),(5,1)\}$$

Note that the elements in the given set are not ordered. It is purposely given this way to emphasize that order is requirement of ordered pair – not that of a set.

Numbers of elements

We have seen that ordered pairs are represented graphically by the points of intersection. The numbers of intersections equal to the product of numbers of rows and columns. Thus, if there are “p” elements in the set “A” and “q” elements in the set “B”, then total numbers of ordered pairs are “pq”. In symbolic notation,

$$n(A \times B) = pq$$

Multiple products

Like other set operations, the product operation can also be applied to a series of sets in sequence. If $A_{1,} A_{2,} \ldots\ldots\ldots, A_n$ is a finite family of sets, then their Cartesian product, one after another, is symbolically represented as :

$$A_1 \times A_2 \times \ldots\ldots\ldots\ldots\ldots\ldots \times A_n$$

This product is set of group of ordered elements. Each group of ordered elements comprises of “n” elements. This is stated as :

$$A_1 \times A_2 \times \ldots \times A_n = \{(x_{1,} x_{2,} \ldots, x_n) : x_1 \in A_{1,} x_2 \in A_{2,} \ldots, x_n \in A_n\}$$

Ordered triplets

The Cartesian product $A \times A \times A$ is set of triplets. This product is defined as :

$$A \times A \times A = \{(x, y, z) : x, y, z \in A\}$$

We can also represent Cartesian product of a given set with itself in terms of Cartesian power. In general,

$$\Rightarrow A^n = A \times A \times \ldots\ldots \times A$$

where “n” is the Cartesian power. If n = 2, then

$$\Rightarrow A^2 = A \times A$$

This Cartesian product is also called Cartesian square.

Example

Problem 3 : If A = {-1,1}, then find Cartesian cube of set A.

Solution : Following the method of writing ordered sequence of numbers, the product can be written as :

$$A \times A \times A = \{(-1,-1,-1),(-1,-1,1),(-1,1,-1),$$
$$(-1,1,1),(1,-1,-1),(1,-1,1),(1,1,-1),(1,1,1)\}$$

The total numbers of elements are 2x2x2 = 8.

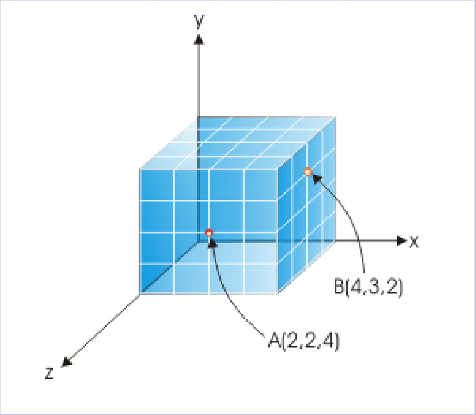
Cartesian Coordinate system

The Cartesian product, consisting of ordered triplets of real numbers, represents Cartesian three dimensional space.

$$R \times R \times R = \{(x,y,z) : x,y,z \in R\}$$

Each of the elements in the ordered triplet is a coordinate along an axis and each ordered triplet denotes a point in three dimensional coordinate space.

Cartesian coordinate system



The coordinate of a point is an ordered tripplet.

Similarly, the Cartesian product " $R \times R$ " consisting of ordered pairs defines a Cartesian plane or Cartesian coordinates of two dimensions. It is for this reason that we call three dimensional rectangular coordinate system as Cartesian coordinate system.

Commutative property of Cartesian product

The Cartesian product is set of ordered pair. Now, the order of elements in the ordered pair depends on the position of sets across product sign. If sets "A" and "B" are unequal and non-empty sets, then :

$$A \times B \neq B \times A$$

In general, any operation involving Cartesian product that changes the "order" in the "ordered pair" will yield different result.

However, if "A" and "B" are non-empty, but equal sets, then the significance of the order in the "ordered pair" is lost. We can use this fact to formulate a law to verify "equality of sets". Hence, if sets "A" and "B" are two non-empty sets and

$$A \times B = B \times A$$

Then,

$$A = B$$

It can also be verified that this condition is true other way also. If sets "A" and "B" are equal sets, then $A \times B = B \times A$. The two way conditional statements can be symbolically represented with the help of two ways arrow,

$$A \times B = B \times A \Leftrightarrow A = B$$

Distributive property of product operator

The distributive property of product operator holds for other set operators like union, intersection and difference operators. We write equations involving distribution of product operator for each of other operators as :

$$A \times (B \cup C) = (A \times B) \cup (A \times C)$$

$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$

$$A \times (B - C) = (A \times B) - (A \times C)$$

Here, sets “A”,”B” and “C” are non-empty sets. In order to ascertain distributive property product operator over other set operators we need to check validity of the equations given above.

We can check these relations proceeding from the defining statements. For the time being, we reason that sequence of operation on either side of the equation does not affect the “order” in the “ordered pair”. Hence, distributive property should hold for product operator over three named operators. Let us check this with an example :

$$A = \{a, b\}, \quad B = \{1,2\} \quad and \quad C = \{2,3\}$$

1: For distribution over union operator

$$\Rightarrow \text{LHS} = A \times (B \cup C) = \{a, b\} \times \{1,2,3\}$$

$$\Rightarrow \text{LHS} = \{(a,1),(a,2),(a,3),(b,1),(b,2),(b,3)\}$$

Similarly,

$$\Rightarrow \text{RHS} = (A \times B) \cup (A \times C) = \{(a,1),(a,2),(b,1),(b,2)\} \cup \{(a,2),(a,3),(b,2),(b,3)\}$$

$$\Rightarrow \text{RHS} = \{(a,1),(a,2),(a,3),(b,1),(b,2),(b,3)\}$$

Hence,

$$\Rightarrow A \times (B \cup C) = (A \times B) \cup (A \times C)$$

2: For distribution over intersection operator

$$\Rightarrow \text{LHS} = A \times (B \cap C) = \{a,b\} \times \{2\}$$

$$\Rightarrow \text{LHS} = \{(a,2),(b,2)\}$$

Similarly,

$$\Rightarrow \text{RHS} = (A \times B) \cap (A \times C) = \{(a,1),(a,2),(b,1),(b,2)\} \cap \{(a,2),(a,3),(b,2),(b,3)\}$$

$$\Rightarrow \text{RHS} = \{(a,2),(b,2)\}$$

Hence,

$$\Rightarrow A \times (B \cap C) = (A \times B) \cap (A \times C)$$

3: For distribution over difference operator

$$\Rightarrow \text{LHS} = A \times (B - C) = \{a,b\} \times \{1\}$$

$$\Rightarrow \text{LHS} = \{(a,1),(b,1)\}$$

Similarly,

$$\Rightarrow \text{RHS} = (A \times B) - (A \times C) = \{(a,1),(a,2),(b,1),(b,2)\} - \{(a,2),(a,3),(b,2),(b,3)\}$$

$$\Rightarrow \text{RHS} = \{(a,1),(b,1)\}$$

Hence,

$$\Rightarrow A \times (B - C) = (A \times B) - (A \times C)$$

Analytical proof

Let us consider an arbitrary ordered pair (x,y), which belongs to Cartesian product set “ $A \times (B \cup C)$ ”. Then,

$$\Rightarrow (x,y) \in A \times (B \cup C)$$

By the definition of product of two sets,

$$\Rightarrow x \in A \quad and \quad y \in (B \cup C)$$

By the definition of union of two sets,

$$\Rightarrow x \in A \quad and \quad (y \in B \quad or \quad y \in C)$$

$$\Rightarrow (x \in A \quad and \quad y \in B) \quad or \quad (x \in A \quad and \quad y \in C)$$

$$\Rightarrow (x, y) \in A \times B \quad or \quad (x, y) \in A \times C$$

By the definition of union of two sets,

$$\Rightarrow (x, y) \in (A \times B) \cup (A \times C)$$

But, we had started with " $A \times (B \cup C)$ " and used definitions to show that ordered pair “(x,y)” belongs to another set. It means that the other set consists of the elements of the first set – at the least. Thus,

$$\Rightarrow A \times (B \cup C) \subset (A \times B) \cup (A \times C)$$

Similarly, we can start with " $(A \times B) \cup (A \times C)$ " and reach the conclusion that :

$$\Rightarrow (A \times B) \cup (A \times C) \subset A \times (B \cup C)$$

If sets are subsets of each other, then they are equal. Hence,

$$\Rightarrow A \times (B \cup C) = (A \times B) \cup (A \times C)$$

Proceeding in the same manner, we can also prove distribution of product operator over intersection and difference operators,

$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$

$$A \times (B - C) = (A \times B) - (A \times C)$$

Cartesian Product (exercise)

Note : The results of some of the questions (3 - 7) are of generic nature. As such, they can also be treated as theorems on Cartesian products.

Worked out exercises

Problem 1 : Cartesian product " $A \times B$ " consists of 6 elements. If three of these are (1,2), (2,3) and (3,3), then find Cartesian product set " $B \times A$ ".

Solution : We need to know two sets “A” and “B” in order to evaluate " $B \times A$ ". First elements of ordered pairs of " $A \times B$ " are elements of set “A”. Hence, “1”,”2” and “3” are the elements of set “A”. On the other hand, second elements of ordered pairs of AXB are elements of set “B”. Hence, “2” and “3” are elements of set “B”.

Now, it is given that there are total 6 elements in the Cartesian product, which is equal to the product of numbers of elements in two sets i.e. 3 X 2. It means that we have identified all elements of sets “A” and “B”.

$$A = \{1,2,3\}$$

$$B = \{2,3\}$$

Following the rule for writing Cartesian product in terms of ordered pairs, we have :

$$\Rightarrow B \times A = \{(1,2), (1,3), (2,2), (2,3), (3,2), (3,3)\}$$

Problem 2 : Two sets are given as : A = {1,2} and B = {3,4}. Find the total numbers of subsets of " $A \times B$ ". Also write power set of AXB in roaster form.

Solution : The total numbers of elements in the Cartesian product " $A \times B$ " is “pq”, where “p” and “q” are the numbers of elements in the individual sets “A” and “B” respectively. The all possible subsets that can be formed including empty set and the product " $A \times B$ " itself is :

$$n = 2^{pq} = 2^{2X_2} = 2^4 = 16$$

Now, the Cartesian product is :

$$\Rightarrow A \times B = \{(1,3), (1,4), (2,3), (2,4)\}$$

The corresponding power set comprises of empty set, 4 sets with elements comprising of one element plus 6 sets with elements comprising of two elements taken at a time plus 4 sets with elements comprising of three elements taken at a time plus set itself. There are total of 16 subsets. The power set is set of all subsets as its elements :

$$\Rightarrow P(A \times B) = \{\varphi, \{(1{,}3)\}, \{(1{,}4)\}, \{(2{,}3)\}, \{(2{,}4)\},$$

$$\{(1{,}3), (1{,}4)\}, \{(1{,}3), (2{,}3)\}, \{(1{,}3), (2{,}4)\}, \{(1{,}4), (2{,}3)\}, \{(1{,}4), (2{,}4)\}, \{(2{,}3), (2{,}4)\},$$

$$\{(1{,}3), (1{,}4), (2{,}3)\}, \{(1{,}3), (1{,}4), (2{,}4)\}, \{(1{,}3), (2{,}3), (2{,}4)\}, \{(1{,}4), (2{,}3), (2{,}4)\}$$

$$\{(1{,}3), (1{,}4), (2{,}3), (2{,}4)\}\}$$

It is easy to follow a scheme to write combination in which order is not relevant. We can denote each of the ordered pair with a symbol like :

$$A \times B = \{a, b, c, d\}$$

As pointed out for generating combination for ordered pair, we can start with the left element and keep changing the last element of the combination till all combinations are exhausted. Here, power set in terms of symbols is :

$$P(A \times B) = \{\varphi, \{a\}, \{b\}, \{c\}, \{d\},$$

$$\{(a, b), \{(a, c)\}, \{(a, d)\}, \{(b, c), \{(b, d)\}, \{(c, d)\},$$

$$\{(a, b, c)\}, \{(a, b, d)\}, \{(a, c, d)\}, \{(b, c, d)\},$$

$$\{(a, b, c, d)\}\}$$

Problem 3 : If " $A \subset B$ " and "C" is any non-empty set, then prove that :

$$A \times C \subset B \times C$$

Solution : Let us first discuss the logic of the relation here. The elements of set “A” are common to set “B”. Now Cartesian product of set "A" with set “C” will yield ordered pairs, which are common with the ordered pairs of the Cartesian product "B" with "C". However, as set “B” is either larger than or equal to, but not smaller than “A”, it follows that above relation should hold.

Now, we prove the relation analytically. Let an arbitrary ordered pair (x,y) belongs to Cartesian product " $A \times C$ ".

$$(x, y) \in A \times C$$

According to definition of Cartesian product,

$$\Rightarrow x \in A \quad and \quad y \in C$$

But “A” is subset of “B”. Hence, $x \in B$.

$$\Rightarrow x \in B \quad and \quad y \in C$$

Again, applying definition of Cartesian product,

$$(x, y) \in B \times C$$

This means that :

$$\Rightarrow A \times C \subset B \times C$$

Problem 4 : If $A \subset B \quad and \quad C \subset D$, then prove that :

$$A \times C \subset B \times D$$

Solution : Let an arbitrary ordered pair (x,y) belongs to Cartesian product " $A \times C$ ".

$$(x, y) \subset A \times C$$

According to definition of Cartesian product,

$$\Rightarrow x \in A \quad and \quad y \in C$$

But “A” is subset of “B”. Hence, $x \in B$. Also, “C” is subset of “D”. Hence, $y \in D$.

$$\Rightarrow x \in B \quad and \quad y \in D$$

Again, applying definition of Cartesian product,

$$\Rightarrow (x, y) \in CXD$$

This means that :

$$\Rightarrow A \times C \subset B \times D$$

Problem 5 : For any given four sets “A”, “B”, “C” and “D”, prove that :

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

Solution : Let an arbitrary ordered pair (x,y) belongs to intersection set “ $(A \times B) \cap (C \times D)$ ”. Then,

$$(x, y) \in (A \times B) \cap (C \times D)$$

Applying definition of intersection,

$$\Rightarrow (x, y) \in A \times B \quad and \quad (x, y) \in C \times D$$

Applying definition of Cartesian product,

$$\Rightarrow (x \in A \quad and \quad y \in B) \quad and \quad (x \in C \quad and \quad y \in D)$$

$$\Rightarrow (x \in A \quad and \quad x \in C) \quad and \quad (y \in B \quad and \quad y \in D)$$

Applying definition of intersection,

$$\Rightarrow (x \in A \cap C) \quad and \quad (y \in B \cap D)$$

Again, applying definition of Cartesian product,

$$\Rightarrow (x, y) \in (A \cap C) X (B \cap D)$$

This means that :

$$\Rightarrow (A \times B) \cap (C \times D) \subset (A \cap C) \times (B \cap D)$$

Similarly, starting from RHS, we can prove that :

$$\Rightarrow (A \cap C) \times (B \cap D) \subset (A \times B) \cap (C \times D)$$

If sets are subsets of each other, then they are equal. Hence,

$$\Rightarrow (A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

Problem 6 : Let “A” and “B” be two non-empty sets. If the numbers of common elements be “n” between sets “A” and “B”, then find the common elements between Cartesian products “ $A \times B$ ” and “ $B \times A$ ”.

Solution : The common elements between sets “A” and “B” is “n”. This means :

$$n(A \cap B) = n$$

We are required to evaluate the expression,

$$n[(A \times B) \cap (B \times A)]$$

We have earlier seen that four given sets,

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

If we put C = B and D = A in this equation, then expression on the left hand side of the equation becomes what is required.

$$\Rightarrow (A \times B) \cap (B \times A) = (A \cap B) \times (B \cap A)$$

$$\Rightarrow n[(A \times B) \cap (B \times A)] = n[(A \cap B)] \times n[(B \cap A)]$$

$$\Rightarrow n\big[(A \times B) \cap (B \times A)\big] = nXn = n^2$$

Problem 7 : Let “A”, “B” and “C” be three sets. Then prove that :

$$AX(B'\cup C')' = (A \times B) \cap (A \times C)$$

Solution : From first De-morgan’s theorem, we know that :

$$(B \cup C)' = B'\cap C'$$

Applying to the LHS, we have :

$$\Rightarrow A \times (B'\cup C')' = A \times [(B')'\cap(C')']$$

Now, component of complement set is set itself. Hence,

$$\Rightarrow A \times (B'\cup C')' = A \times (B \cap C)$$

Applying distributive property of product operator over intersection operator, we have :

$$\Rightarrow A \times (B'\cup C')' = (A \times B) \cap (A \times C)$$

Problem 8 : Let “A”, “B” and “C” be three sets. Then prove that :

$$AX(B'\cap C')' = (A \times B) \cup (A \times C)$$

Solution : From second De-morgan’s theorem, we know that :

$$(B \cap C)' = B'\cup C'$$

Applying to the LHS, we have :

$$\Rightarrow A \times (B'\cap C')' = A \times [(B')'\cup(C')']$$

Now, component of complement set is set itself. Hence,

$$\Rightarrow A \times (B'\cap C')' = A \times (B \cup C)$$

Applying distributive property of product operator over union operator, we have :

$$\Rightarrow A \times (B' \cap C')' = (A \times B) \cup (A \times C)$$

Relations

We encounter different types of relationship in our daily life. Besides human relationship that we are so familiar with, there are numerous other relationships, including those which are purely abstraction of processes and ideas.

In essence, any two elements which are paired have potential to possess relationship between them. Now think of the Cartesian product that we have defined for two sets. It consists of ordered pairs of elements - one each from the two sets. The total numbers of ordered pairs in a Cartesian product is equal to the product of numbers of elements in each set. A particular relation, however, may not comprise all ordered pairs.

In this module, we shall limit our discussion to binary relations only. A binary relation is a relation as defined between two elements either from the same set of from two different sets.

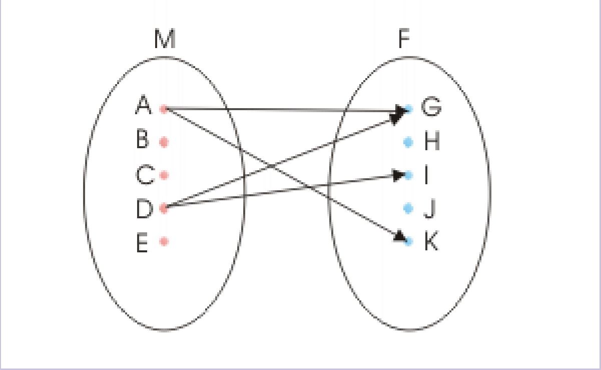
Consider a “get together”. Divide the people in two groups comprising of males and females. A certain numbers of ordered pairs will qualify for a relation say “classmate of” – not all. Similarly, a relation such as “brother of” may include some of the ordered pairs of the Cartesian product of two set.

$$M = \{A, B, C, D, E\}$$

$$F = \{G, H, I, J, K\}$$

We can represent the relationship of “classmate of” as shown here :

Relation



Relation between elements of two sets

From the figure, we can write the collection of relationship “classmate of” as a set of ordered pairs of two sets,

$$R(\text{classmate of}) = \{(A, G), (A, K), (D, G), (D, I)\}$$

In the nutshell, we can think of relation as a collection (set), which comprises of ordered pairs (instances of relation). Note that it is a specific relation. This is a relation between elements of two sets. Clearly, this relation set can not exceed the Cartesian product of two sets under consideration. Thus, a relation set is a subset of Cartesian product set.

Relation

> A relation “R” from a non-empty set “A” to non-empty set “B” is a subset of Cartesian product “AXB”.

We need to note that a relation is defined in a particular order “from” to “to”. It is for this reason, we denoted relation pictorially by an arrow which is directed from the elements of "from" set “A” to "to" set “B”.

Consider few examples,

$$R = \left\{(x, y) : x = y^2, x \in A, y \in B\right\}$$

$$R = \{(x, y) : x = y + 1, x \in A, y \in B\}$$

In certain circumstance, we are required to work with relation among the elements of the same set. For example, consider the male set defined earlier. Some of the male members may as well be classmates and hence related to each other. Such relation is relation on one set only and is called "relation on A" or "relation on B" etc. Few examples are :

$$R = \{(x, y) : x = x + 1, x \in A\}$$

$$R = \{(x, y) : y = x + 1 \quad and \quad x, y \in A\}$$

Example

Problem 1 : Let $A = \{1,2, \ldots\ldots, 10\}$. Write down the relation set in roaster form, which is defined as :

$$R = \{(x, y) : y = 3x \quad and \quad x, y \in A\}$$

Solution : We begin with the first element of “A” i.e. x =1. Since other element also belongs to set “A”, it is required that the value of “y” be one of the elements in the set "A".

$$For \quad x = 1, \quad y = 3X1 = 3$$

$$For \quad x = 2, \quad y = 3x = 3X2 = 6$$

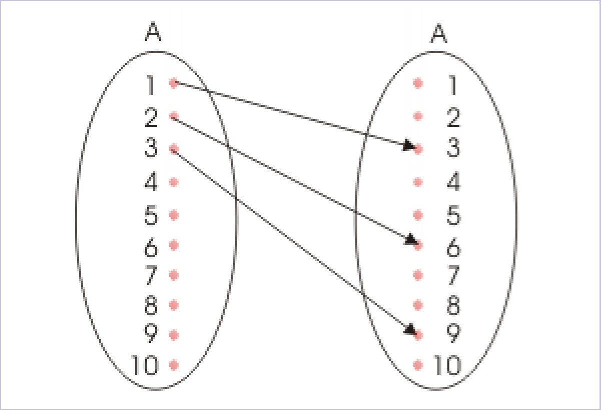
$$For \quad x = 3, \quad y = 3x = 3X3 = 9$$

Thus, “x” can assume values “1”,”2” and “3” for which “y” can assume values “3”,”6” and “9” respectively in accordance with the given relation. The relation, therefore, is set of ordered pairs :

$$R = \{(1,3), (2,6), (3,9)\}$$

We can visualize the relation pictorially as shown in the figure.

Relation on self



Relation between elements of

the same set

Domain of the relation

The domain represents the valid values of the first element of the ordered pairs in the relation. Clearly, the elements of domain of a relation belong to “from” set “A”. But, elements in the domain are only those which are valid for the relation. It means that domain does not consist of all elements of “from” set “A”. Thus, domain set is a subset of “from” set “A”.

Domain

> The set of first elements of all ordered pairs in the relation “R” from set “A” to “B” is called the domain of relation “R”.

We can write the domain set of relation “R” from set “A” to set “B” in set builder form as :

$$\text{Domain}(R) = \{x : (x, y) \in R\}$$

Consider the example given earlier. The relation set is :

$$R = \{(1{,}3), (2{,}6), (3{,}9)\}$$

The domain according to definition is :

$$\text{Domain}(R) = \{1{,}2{,}3\}$$

$$\Rightarrow \text{Domain}(R) \subset A$$

Co-domain

In a relation “R” from set “A” to “B”, the set “B” is called co-domain.

Range of the relation

The range represents the valid values of the second element of the ordered pairs in the relation.

Range

> The set of second elements of all ordered pairs in the relation “R” from set “A” to “B” is called the range of relation “R”.

We can write the range set of relation “R” from set “A” to set “B” in set builder form as :

$$\text{Range}(R) = \{y : (x, y) \in R\}$$

Consider the example given earlier. The relation set is :

$$R = \{(1,3), (2,6), (3,9)\}$$

The range according to definition is :

$$\text{Range}(R) = \{3,6,9\}$$

Clearly, the elements of range of a relation belong to “to” set “B”. But, elements in the range are only those which are valid for the relation. It means that range does not consist of all elements of “to” set “B”. Thus, range set is a subset of “to” set “B”.

$$\text{Range}(R) \subset B$$

Example

Problem 2 : Let $A = \{1,2,3,4,5\}$ and $B = \{1,4,5\}$. Let a relation from “A” to “B” is :

$$R = \{(x, y) : x < y, \quad (x, y) \in AXB\}$$

Find R, Domain (R) and Range(R).

Solution : Let us find “y” for each value of “x”.

$$\text{For} \quad x = 1, \quad y = 4,5$$

$$For \quad x = 2, \quad y = 4{,}5$$

$$For \quad x = 3, \quad y = 4{,}5$$

$$For \quad x = 4, \quad y = 5$$

$$For \quad x = 5, \quad \text{There is no value of “y” in set “B”}$$

Hence,

$$R = \{(1{,}4), (1{,}5), (2{,}4), (2{,}5), (3{,}4), (3{,}5), (4{,}5)\}$$

$$\text{Domain}(R) = \{1{,}2{,}3{,}4\}$$

$$\text{Range}(R) = \{4{,}5\}$$

Numbers of relations

Between two sets, the Cartesian product set consists of all possible instances of relation as ordered pair. Here, we need to find the total possible relations that can be generated from these ordered pairs. We have seen that total numbers of ordered pairs in the Cartesian product of sets “A” and “B” is “pq”, where “p” and “q” are the numbers of elements in two sets respectively.

Now, relation is nothing but a subset of the Cartesian product. It means that total numbers of relation is equal to total numbers of possible subsets of the Cartesian product. Recall that the set formed from all possible subsets is called power set. The numbers of subsets in the power set is given by

$$n = 2^{pq}$$

Clearly, this number “n” denotes all possible relations (subsets) that can be generated from two finite sets. We should, however, be careful in interpreting this number as it also contains the mandatory empty set, which is not a meaningful set from the point of view of relation.

Inverse relation

Inverse relation of a given relation “R” from set “A” to set “B” is set of ordered pairs in which first and second elements exchanges their positions. The inverse set is defined in reference to a given relation. The inverse relation of a given relation “R” from “A” to “B” is denoted as “ R^{-1}”. Clearly,

$$R^{-1} = \left\{(y, x) : (x, y) \in R\right\}$$

where,

$$R = \{(x, y) : (x, y) \in AXB\}$$

We should be careful to understand that “-1” is not a power, but a part of symbol to represent inverse relation with respect to a given relation. It is also clear that :

$$If (x, y) \in R \Leftrightarrow (y, x) \in R^{-1}.$$

As the elements in the ordered pair of the relation exchanges positions, domain and range sets are exchanged across the sets.

$$\text{Domain}\left(R^{-1}\right) = \text{Range}(R)$$

$$\text{Range}\left(R^{-1}\right) = \text{Domain}(R)$$

Example

Problem 3 : Let $A = \{1,2, \ldots\ldots, 10\}$. A relation on “A” is defined as

$$R = \{(x, y) : y = 3x, \quad where \quad x, y \in A\}$$

Find R^{-1} , Domain (R^{-1}) and Range (R^{-1}).

Solution : In the earlier example, we determined the relation “R” as :

$$R = \{(1,3), (2,6), (3,9)\}$$

According to the definition of inverse set, the elements of the ordered pair in the relation set are exchanged :

$$R^{-1} = \{(3,1), (6,2), (9,3)\}$$

Clearly, inverse relation can be represented in set builder form as :

$$\Rightarrow R^{-1} = \{(y, x) : y = x/3, \quad where \quad x, y \in A\}$$

Now, the domain and range of R^{-1} are :

$$Domain(R^{-1}) = \{3,6,2\}$$

$$\text{Range}(R^{-1}) = (1,2,3)$$

Relation types

Relations among elements of a set have wide possibilities. A systematic approach to study them is facilitated by recognizing different relation types. It should be noted that all relation types described here are relation on one set.

We describe a relation on set itself as :

Relation on A
A relation “R” from set “A” to “A” is called a “relation on A”.

In this module, we shall be using a symbol, “xRy” to denote an instance of relation (ordered pair). The symbol conveys that the instance of relation denoted by the symbol is an ordered pair (x,y), which follows relation “R”.

Void relation

Relation is a subset of Cartesian product of two sets. We have seen that power set of Cartesian product “ $A \times B$ ” is a set of all possible relations among the elements of sets “A” and “B”. In the case of “relation on A”, the power set of Cartesian product “ $A \times A$ ” is a set of all possible relations among the elements of set “A”.

One of the subsets of the power set is empty set or void set. This subset without any element is called the void relation.

$$R = \varphi = \{\}$$

Universal relation

Universal relation is the widest possible relation. This relation consists of all ordered pairs of the Cartesian product “ $A \times A$ ”.

$$R = A \times A$$

Consider a set $A = \{1,2,3\}$. Then, universal relation set is :

$$R = \{(1{,}1), (1{,}2), (1{,}3), (2{,}1),$$

$$(2{,}2), (2{,}3), (3{,}1), (3{,}2), (3{,}3)\}$$

Identity relation

An identity relation is defined as :

Identity relation

> In an identity relation "R", every element of the set “A” is related to itself only.

Note the conditions conveyed through words “every” and “only”. The word “every” conveys that identity relation consists of ordered pairs of element with itself - all of them. The word “only” conveys that this relation does not consist of any other combination.

Consider a set $A = \{1{,}2{,}3\}$. Then, its identity relation is :

$$R = \{(1{,}1), (2{,}2), (3{,}3)\}$$

It is evident that a set has only one such relation. This relation, as we can see, identifies the set - as it identifies each elements of the set, which are related to itself. By looking at the relation, we can identify the set itself. For this reason, the name of this relation is identity relation. In set builder form, we express an identity relation as

$$R = \{(x, x) : \quad \text{for all} \quad x \in A\}$$

The qualification of the relation is that first and second element of the ordered pair is same element, which belongs to set A.

The followings are not an identity relation :

$$R_1 = \{(1{,}1), (2{,}2)\}$$

$$R_2 = \{(1{,}1), (2{,}2), (3{,}3), (1{,}2), (1{,}3)\}$$

First one is not an identity relation as it does not include the pairing of remaining element “3”. Second is not an identity relation, because there are other combinations of pairs in the relation.

Reflexive relation

Reflexive relation is an expansion of identity relation. In the simple word, reflexive relation is plus identity relation.

Reflexive relation

> In reflexive relation, "R", every element of the set “A” is related to itself.

The definition of reflexive relation is exactly same as that of identity relation except that it misses the word “only” in the end of the sentence. The implication is that this relation includes identity relation and permits other combination of paired elements as well.

Consider a set $A = \{1,2,3\}$. Then, one of the possible reflexive relations can be :

$$R = \{(1,1), (2,2), (3,3), (1,2), (1,3)\}$$

However, following is not a reflexive relation :

$$R_1 = \{(1,1), (2,2), (1,2), (1,3)\}$$

It is not a reflexive relation as one instance of identity relation (3,3) is absent and violates the condition that every element of the set is related to itself.

We state the condition for reflexive relation as :

$$R \quad \text{is reflexive} \quad \Leftrightarrow (x, x) \in R, \quad \text{for all} \quad x \in A$$

It is clear that identity relation is a reflexive relation. Further, universal relation consists of all combinations of ordered pairs in the Cartesian product. It means it consists of all elements of the identity relation apart from other ordered pairs. Hence, universal relation is also a reflexive relation.

Interpretation of reflexive relation

Reflexivity of a relation (meaning that a relation is reflexive) is used to characterize important algebraic relations. Following relations are reflexive :

- “is equal to”
- “is less than or equal to”
- “is greater than or equal to”
- “divides”
- “is subset of”

The relation “is less than” or “greater than”, however, are not reflexive.

Examples

Problem 1 : Determine whether “greater than or equal to” is a reflexive relation for natural number.

Solution : A relation, “R”, representing “greater than or equal to” is defined as relation on natural number (N) as :

$$(x, y) \in R \Leftrightarrow x \geq y \quad \text{where} \quad x, y \in N$$

We construct data for “x” and “y” in accordance with the given relation for few initial natural numbers, say 1, 2 and 3, as under :

$$\text{For} \quad x = 1, \quad y = 1,2,3$$

$$For \quad x = 2, \quad y = 2,3$$

$$For \quad x = 3, \quad y = 3$$

Thus, the relation set is :

$$R = \{(1,1), (1,2), (1,3), (2,2), (2,3), (3,3)\}$$

Evidently, this set consists of relation of all elements of the set, which are related to itself ie. (1,1), (2,2) and (3,3). Thus, we conclude that “is greater than or equal to” is a reflexive relation.

Problem 2 : Determine whether “is not equal to” is a reflexive relation for natural number?

Solution : A relation, “R”, representing “is not equal to” is defined as relation on natural number (N) as :

$$(x,y) \in R \Leftrightarrow x \neq y \quad where \quad x,y \in N$$

We construct data for “x” and “y” in accordance with the given relation for few initial natural numbers, say 1,2 and 3, as under :

$$For \quad x = 1, \quad y = 2,3$$

$$For \quad x = 2, \quad y = 1,3$$

$$For \quad x = 3, \quad y = 1,2$$

Thus, the relation set is :

$$R = \{(1,2), (1,3), (2,1), (2,3), (3,1), (3,2)\}$$

Evidently, this set does consists of all ordered pair representing relation of an element with itself. The instances (1,1), (2,2) and (3,3) are missing. Thus, we conclude that “is not equal to" is a irreflexive relation.

Symmetric relation

In symmetric relation, the instance of relation has a mirror image. It means that if (1,3) is an instance, then (3,1) is also an instance in the relation. Clearly, an ordered pair of element with itself like (1,1) or (2,2) are themselves their mirror images. Consider some of the examples of the symmetric relation,

$$R_1 = \{(1,2), (2,1), (1,3), (3,1)\}$$

$$R_2 = \{(1,2), (1,3), (2,1), (3,1), (3,3)\}$$

We have purposely jumbled up ordered pairs to emphasize that order of elements in relation is not important. In order to decide symmetry of a relation, we need to identify mirror pairs. We state the condition of symmetric relation as :

$$Iff \quad (x,y) \in R \Rightarrow (y,x) \in R \quad \text{for all} \quad x,y \in A$$

The symbol “Iff” means “If and only if”. Here one directional arrow means “implies”. Alternatively, the condition of symmetric relation can be stated as :

$$xRy \Rightarrow yRx \quad \text{for all} \quad x,y \in A$$

In words, we say that if (x,y) be an instance of relation, then (y,x) will also be the instance of a symmetric relation "R".

It is clear that identity relation is a symmetric relation. Also, universal set consists of the Cartesian product of a set with itself. It means that the relation consists of instances with mirror instances. Therefore, universal relation is also symmetric relation.

Symmetric and inverse relation

An inverse relation (R^{-1}) consists of ordered pairs with exchange of positions of the elements in a given relation (R). Now let us consider a symmetric relation,

$$R = \{(1,2),(1,3),(2,1),(3,1),(3,3)\}$$

By definition, its inverse relation is :

$$\Rightarrow R^{-1} = \left\{(2,1),(3,1),(1,2),(1,3),(3,3)\right\}$$

Using the fact that order does not change a set, we conclude that :

$$\Rightarrow R = R^{-1}$$

We use this fact to identify symmetric relation. The given set is a symmetric relation, if it equals its inverse set.

Analytical proof

Let “R” be a symmetric relation on set “A”. In order to prove that $R = R^{-1}$, we consider an arbitrary instance of relation “R” :

$$(x,y) \in R$$

According to definition of symmetric relation,

$$(y,x) \in R$$

According to definition of inverse relation,

$$(x,y) \in R^{-1}$$

But, we had started with “R” and used definitions to show that “(x,y)” belongs to another set “R^{-1}”. It means that the “R^{-1}”set consists of the elements of set “R” – at the least. Thus,

$$R \subset R^{-1}$$

Similarly, we can start with “R^{-1}”set and reach the conclusion that :

$R^{-1} \subset R$

If sets are subsets of each other, then they are equal. Hence,

$$\Rightarrow R = R^{-1}$$

Asymmetric relation

A relation “R” on a set “A” is asymmetric for the following condition :

$$Iff \quad (x, y) \in R \quad \text{and} \quad (y, x) \Rightarrow a = b \quad \text{for all} \quad x, y \in A$$

It means that possibility of symmetry in asymmetric relation exists only if elements are equal.

Transitive relation

If “R” be the relation on set A, then we state the condition of transitive relation as :

$$Iff \quad (x, y) \in R \quad \text{and} \quad (y, z) \in R \Rightarrow (x, z) \in R \quad \text{for all} \quad a, b, c \in A$$

Alternatively,

$$xRy \quad and \quad yRz \Rightarrow xRz \quad \text{for all} \quad x, y, z \in A$$

In words, we say that if (x,y) and (y,z) be the instances of a relation R such that (a,z) is also the instance of the relation, then that relation is transitive.

The identity and universal relations are transitive. Some other important transitive relations are :

- “is equal to”
- “is greater than”
- “is at least as great as”
- “is a subset of”
- "divides"

Example

Problem 3 : Determine whether “divides” is a transitive relation for natural number?

Solution : Let us consider three elements “x”,”y” and “z” of set “N” of natural numbers such that a relation “R” on “N” is :

$$(x, y) \in R, (y, z) \in R, \quad \text{``divides''} \quad , x, y, z \in N$$

This means that :

$$\text{``x divides y''} \quad \text{and} \quad \text{``y divides z''}$$

Let us now consider two natural numbers “a” and “b” such that :

$$y = ax \quad \text{and} \quad z = by$$

$$z = abx$$

This means that “x divides z”. Hence, we conclude that the relation "divides" is transitive relation.

Equivalence relation

A relation is equivalence relation if it is reflexive, symmetric and transitive at the same time. In order to check whether a relation is equivalent or not, we need to check all three characterizations.

Functions

Function is a special relation. It is also conceived as a “rule”, because function is a relation between elements of two sets, following certain rule. Every element of a set (say A) is related to exactly one element of other set (say B). This relationship is described as mapping of all elements of one set to elements of another set.

In order to emphasize, we need to enumerate the way “function” relation is special :

1. Every element of set “A” is related to elements in set “B”.
2. An element of set “A” is related to exactly one element of “B”.

It can be deduced from the above characterization of a function that the element of set “B” may be paired with none or one or more elements of set “A”.

In order to illustrate function relation, let us consider an example. Let "A" and "B" be two sets as given here :

$$A = \{-1,0,1,2,3\}$$

$$B = \{-1,0,1,2,3,4,5,6,7,8,9,10\}$$

The two sets are related by the relation :

$$R = \left\{(x, y) : y = x^2 - 1, x \in A, y \in B\right\}$$

The values of “y” for given values of “x” are :

$$For \quad x = -1, \quad y = 1 - 1 = 0$$

$$For \quad x = 0, \quad y = 0 - 1 = -1$$

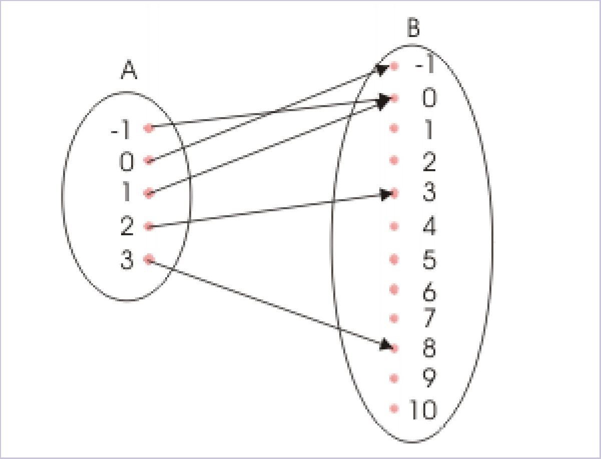
$$For \quad x = 1, \quad y = 1 - 1 = 0$$

$$For \quad x = 2, \quad y = 4 - 1 = 3$$

$$For \quad x = 3, \quad y = 9 - 1 = 8$$

The relation between two sets is pictorially shown with arrow diagram. We note that all elements of “A” are mapped. Further, elements in “A” are uniquely mapped i.e. they are paired to exactly one element of set “B”. It is, however, possible that few of the elements in set “A” are related to same element in “B” like "-1" and "1" in set "A" are related to "0" in set “B”. In the nutshell, we see that this relation meets both properties as enumerated for a function relation and hence is a function relation.

Function relation



Every element of set “A” is related to exactly one element in set “B”.

Looking in reverse direction, we see that elements in “B” may be paired – with no element (1,2,4,5,6,7,9,10) or with one element (-1,3,8) or with more than one element (0) in "A".

We generally drop word “relation” from “function relation” and call it simply as “function”. The function is denoted by a small letter like “f”. To elaborate the direction of function, we expand the symbol as :

$$f : A \rightarrow B$$

This means that function is mapped from “A” to “B”. Now, in order to define the function, we need to understand the concept of “image” and “pre-image” elements. We call first element “x” of set “A” in the ordered pair (x,y) of the function as the “pre-image” of second element “y” of set “B”. The second element “y” of set “B” is called the “image” of the first element “x” of set “A”.

The image is also denoted as “f(x)”. We read “f(x)” as image of “x” under rule “f”. For a particular value of x = a, "f(a)" is a particular instance of image :

$$f(a) = b$$

Function

> A relation “f” is a function, if every element in set “A” has one and only one image in set “B”.

Domain, range and co-domain of function

As all the elements of set “A” are involved, it emerges that that the set of first elements in the ordered pairs i.e. domain set is same as set “A”. We can not say the same for set "B". The set “B” may have other elements than those, which have been mapped with the elements of “A”. The range is simply the set of images of the function. However, as defined earlier, the set “B” is co-domain of the relation and hence that of function in this special case. It is clear that range is a subset of co-domain "B".

$$\text{Domain of ”f”} = A$$

$$\text{Co-domain of ”f”} = B$$

$$\text{Range of ”f”} = \text{Set of images} = \{f(x) : x \in A\}$$

Equal functions

Two functions are equal, if each ordered pair in one of the two functions is uniquely present in other function. It means that if “g” and “h” be two equal functions, then :

$$g(x) = h(x) \quad \text{for all} \quad ``x"$$

Two functions g(x) and h(x) are equal or identical, if all images of two functions are equal. Further, we can visualize equality of two functions in a negative context. If there exists “x” such that g(x)≠ h(x), then two functions are not equal. We state this symbolically as :

$$\text{If} \quad g(x) \neq h(x), \quad \text{then} \quad f \neq h$$

The important question, however, is that whether equality of functions in terms of equality of images is a sufficient condition? We can see here that two functions can meet the stated condition even if they are constituted by different sets of ordered pair. There may be additional ordered pairs, which are present in one, but not in other.

In order to remove such possibilities, two equal functions should have same domain. This will ensure that set of ordered pairs in two functions are same. We conclude this discussion by saying that two functions are equal, iff

1. g(x) = h(x) for all “x”
2. Domain of “f” = Domain of “h”

It is clear that equality of functions, however, do not require that co-domains be equal.

Real function

If the range of a function is a set of real numbers, then the function is called “real valued function”. In other words, if the range of a function is either the set “R” or its subset, then it is a real valued function. We should emphasize here that “R” denotes set of real number and it is not the symbol for relation, which is also denoted as “R”.

Further, we distinguish “real valued function” from “real function”. The very terminology is indicative of the difference. The term “real valued function” means that the value of function i.e. image is real. It does not say anything about “pre-image”. Now, there can be a function, which accepts non-real complex numbers, but maps to a real value.

On the other hand, a real function has both image and pre-image as real numbers. It follows then that the domain of a “real function” is also either a set or subset of real numbers.

Real function

A function is a real function, if its domain and range are either “R” or subset of “R”.

Note: Our discussion from this point onwards in the course relates to real function only – unless otherwise stated.

Interpretation of function relation

It is intuitive to find similarity of an algebraic equation to the “rule” of a function. Consider an equation,

$$y = x^2 + 1$$

This equation is valid for all real values of “x”. The set of real values of “x”, belongs to set “R”. The set of values of “y” also belongs to set of “R”. On the other hand, the equation itself is the rule that maps two sets comprising of values of “x” and “y”.

Alternatively, we can write the rule also as :

$$\Rightarrow f(x) = x^2 + 1$$

In terms of rule, we define function, saying that :

$$f : R \rightarrow R \quad by \quad f(x) = x^2 + 1$$

We read it as : “f” is a function from “R” to “R” by the rule given by $f(x) = x^2 + 1$.

From this description, we think a function as a relation, which is governed by a specified rule. The rule relates two sets known as domain and co-domain, which are sets of real numbers. One of the quantities “x” is independent of other quantity “y”. The other quantity “y” is a dependent on quantity “x”. In plain words, one of the interpretations is that function relates dependent and independent variables. As a matter of fact, we would attach additional meanings to the concept of function as we proceed to study it in details.

Example

Problem 1 : Let “A” be the set of first three natural numbers. A real function is defined as :

$$f : A \rightarrow N \quad by \quad f(x) = x^2 + 1$$

Find (i) domain of “f” (ii) range of “f” (iii) co-domain of “f” (iv) f(3) and (v) pre-images of 2 and 4.

Solution : Here set “A” is the domain of “f”. Hence,

$$\text{Domain of ”f”} = A = \{1,2,3\}$$

For determining the range, we need to find images for the each element of the domain as :

$$\text{For} \quad x = 1, \quad f(x) = x^2 + 1 = 1^2 + 1 = 2$$

$$\text{For} \quad x = 2, \quad f(x) = x^2 + 1 = 2^2 + 1 = 5$$

$$\text{For} \quad x = 3, \quad f(x) = x^2 + 1 = 3^2 + 1 = 10$$

Hence, range of function is given as :

$$\text{Range of ”f”} = A = \{2,5,10\}$$

Co-domain of the function is the second set on which the elements of first set are mapped. It is given that “f” is a function from set “A” to set “N”.

Hence,

$$\text{Co-domain of "f"} = N$$

The image of set for x = 3 has been already calculated. It is :

$$\Rightarrow f(3) = x^2 + 1 = 3^2 + 1 = 10$$

For pre-image of f(x) = 2, we have :

$$\Rightarrow f(x) = 2 = x^2 + 1$$

$$\Rightarrow x^2 = 1$$

$$\Rightarrow x = 1, -1$$

But, only "1" is an element of domain set "A" - not "-1". Hence, "pre-image" of "2" is "1".

For pre-image of f(x) = 4, we have :

$$\Rightarrow f(x) = 4 = x^2 + 1$$

$$\Rightarrow x^2 = 3$$

$$x = \sqrt{3,} - \sqrt{3}$$

But it is given that domain is first three natural numbers only. Thus, we conclude that “4” has no pre-image.

Numbers of functions

We can find out maximum numbers or total possible numbers of functions that can be generated by the rule from given domain and co-domain sets, provided these sets are finite sets. We have noted that the total numbers of relations generated from Cartesian product of two sets “A” and “B” is given by :

$$N = 2^{pq}$$

where “p” and “q” are the finite numbers of elements in sets “A” and “B”.

However, function is a special relation, in which each element of set ”A” is related to exactly one element of set “B” - unlike in the case of power set in which we count all possible combinations. Hence, number of possible relations is not same as the numbers of possible functions.

For determining total numbers of functions from two given sets, let us consider that “m” and “n” denote the numbers of elements in domain “A” and co-domain “B” respectively. Then, an element of domain can combine with one of the “n” elements in “B”. Hence, total numbers of such relations for a total of “m” elements in set “A” is :

$$N_f = mXn = mn$$

Finite and infinite functions

The numbers of elements of ordered pair in the function set is equal to the numbers of elements in domain set. This follows from the fact that every element of domain set “A” is related to an unique element in “B”. Thus, if domain is a finite set, then the resulting function is finite. Consider the earlier example, when A = {1,2,3} and function is defined as :

$$f : A \rightarrow N \quad by \quad f(x) = x^2 + 1$$

The function set is a finite set :

$$f = \{(1{,}2), (2{,}5), (3{,}10)\}$$

On the other hand, if we expand this function by defining the relation from the infinite set of natural numbers, “N” to “N”, then resulting set of ordered pair is an infinite set and so is the function :

$$g : N \rightarrow N \quad by \quad f(x) = x^2 + 1$$

The resulting function set, in the set builder form, is given as :

$$g = \left\{(x, y) : y = x^2 + 1, \quad where \quad x, y \in N\right\}$$

Function graphs

Here, we shall introduce an alternative way to represent a function. We should be aware that we can define a function even with a graph. Graphical representation of function is intuitive and revealing about their characteristics.

Function is a set of ordered pairs between “x” and “y” from domain and co-domain sets respectively in accordance with certain rule. If we look closely at the function set, then it is easy to realize that the elements of ordered pair (x,y) can be considered to be coordinates “x” and “y” of a planar coordinate system.

We represent independent variable, “x” i.e. the element of domain set “A” as abscissa along x-axis and dependent variable, “y”, i.e. the element of co-domain “B” as ordinate along y-axis. A point on the plot represented by coordinate (x,y) is an instance of or value of the function for a given value of “x”. Compositely, the graph itself is the collection of all such points, which form part of the function set.

For example, we draw a graph, which is defined as :

$$f : N \rightarrow N \quad by \quad f(x) = x, \quad where \quad x \in N$$

In order to plot the function, we evaluate function values for values of “x” :

$$For \quad x = 1, \quad y = 1$$

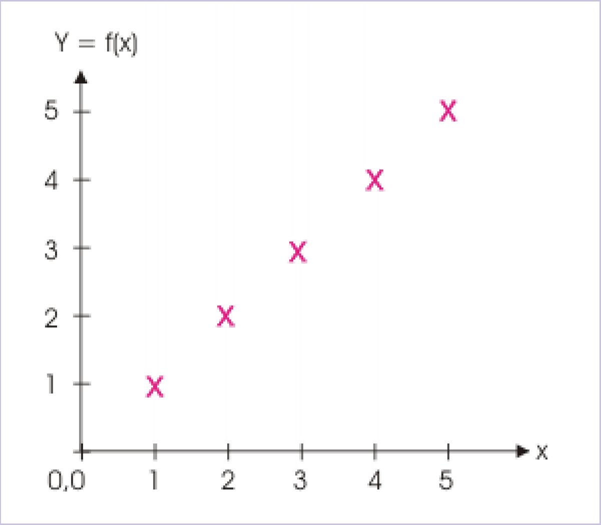
$$For \quad x = 2, \quad y = 2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$For \quad x = n, \quad y = n$$

Function graph



The plot is a collection of discrete points.

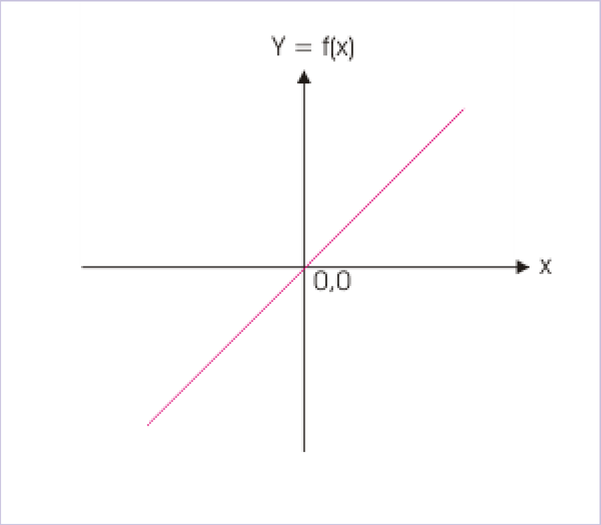
Note that plot of the function is a collection of discrete points only.

For the plot to be continuous, it is clear that the domain and co-domain of the function should be set of real numbers. In that case, we can define the function as :

$$g : R \to R \quad by \quad f(x) = x, \quad where \quad x \in R$$

The corresponding plot is a bisector straight line, passing through the origin, as shown in the figure here :

Function graph



The plot is a continuous straight line passing through origin.

Classification of real functions

Real functions can be classified from different point of views. Here, we present few major classifications.

Based on expression types

1: Algebraic function : The function (function rule) consists of algebraic expression, consisting of terms, which are constituted of constants and variables. The variables of algebraic expressions may be raised to a constant exponent. Example :

$$\frac{\sqrt{(3x^2+2)}}{x}; \quad x \neq 0$$

2: Transcendental function : The non algebraic functions are called transcendental functions. They include logarithmic, exponential, trigonometric and inverse trigonometric functions etc. Example :

$$\sin x + \cos x$$

Based on independent and dependent variables

1: Explicit function : A function is an explicit function, if its dependent variable can be expressed in terms of independent variables only. Example :

$$y = x^2 + 1$$

2: Implicit function : A function is an implicit function, if its dependent variable can not be expressed in terms of independent variables only. Example :

$$xy = \sin(x + y)$$

Function types

The relation between two sets under a rule has two perspectives. We can look at the relation in the direction from domain set “A” to co-domain set “B”. This is the function view. But, we can also look this relation in opposite direction from “B” to “A”.

When we see function relation from domain “A” to co-domain “B”, then we find following possibilities :

- Every element of domain is related to different element of co-domain (**one to one function or injection**)
- More than one elements of domain is related to an element of co-domain (**many to one function)**

When we see relation from co-domain “B” to domain “A”, then we find following possibilities :

- There are elements in co-domain, which are not related to any of the elements in domain (**into function**).
- There are no elements in co-domain, which are not related to elements in domain (**onto function or surjection**).
- There are elements in co-domain, which are related to exactly one element in domain. This statement is an equivalent statement (**one to one function**).
- There are elements in co-domain, which are related to more than one element in domain. This statement is an equivalent statement (**many to one function**).

Thus, we see that there are many ways in which a function - as a relation - can be unique and hence different. This gives rise to function types, which – as we shall see – are reflection of different possibilities that we have enumerated above.

One - one function (Injection)

As is evident, this function describes relation in which something can be related distinctly to something. The countries have unique and distinct capital. It is evident that a function, based on this relation, would be an injection.

One - one function (Injection)

A function $f : A \rightarrow B$ is an injection, if different elements of domain set "A" have different images in co-domain set "B".

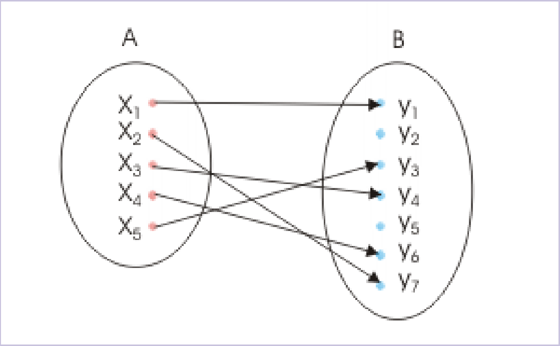
In plain words, every "x" in "A" associates with a distinct "y" in "B". We can yet put it like this : Every argument (x) is related to distinct value (y).

In order to represent the condition of injectivity symbolically, we can think of two different elements "x" and "y" in set "A". Then, two images f(x) and f(y) in "B", corresponding to these elements in "A" are not equal. We capture this intent in constructing condition for an injection as :

$$f : A \rightarrow B \quad \text{is an injection} \quad \Leftrightarrow \quad x \neq y, \quad f(x) \neq f(y) \quad \text{for all} \quad x, y \in A$$

We can also interpret injection by asserting that if two images are equal, then it means that they are images of the same pre-image. The map diagram, corresponding to an injection, is shown in the figure. Note that elements in "A" are mapped to different elements in "B".

one-one function (injection)



Every argument (x) is related to distinct value (y).

Example

Problem 1: Consider a function defined as :

$$f : Z \rightarrow Z \quad by \quad f(x) = x^2 + 1 \quad \text{for all} \quad x \in Z$$

Determine whether the function is an injection?

Solution : We consider two arbitrary elements of the domain set such that :

$$f(x) = f(y)$$

We have deliberately considered a contradictory supposition with respect to the requirement of injectivity. If this supposition yields x = y, then the given function is an injection; otherwise not. Here,

$$\Rightarrow x^2 + 1 = y^2 + 1$$

$$\Rightarrow x^2 = y^2$$

$$\Rightarrow x = \pm y$$

This is not an unique solution. Here, “x” is not uniquely equal to “y”. We conclude that given function is not an injection. As a matter of fact, we can infer a check on our conclusion as,

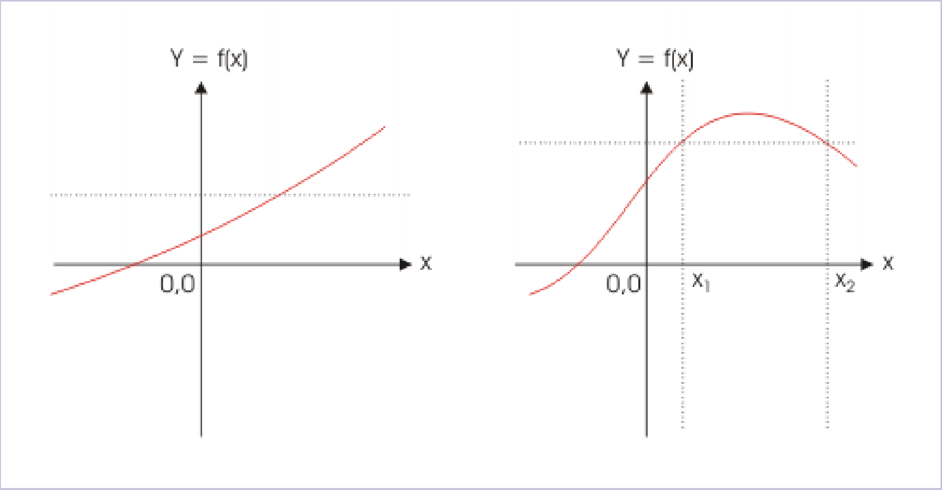
$$\Rightarrow f(1) = f(-1) = 2$$

Thus, we see that two pre-images relate to one image, which is contradictory to the requirement of an injection.

Increasing and decreasing functions

The fact that function value is different for different arguments has an important bearing on the nature of injection plot. Consider two plots shown in the figure. In the plot shown on the left, a straight line parallel to x-axis intersects the plot only once. In the second plot, a line parallel to x-axis intersects the plot at two points for $x = x_1$ and $x = x_2$. The function represented by second plot is not an injection as two values of arguments map to a single value of function – not two different values as required for an injection function.

one-one function (injection)



Injection graph is either increasing or decreasing.

It means that intersection plot intersects a line parallel to x-axis only once. This is possible only if the function is either (i) continuously increasing or (ii) continuously decreasing. Note the use of word “continuously”. An increasing plot, for example, if drops, then we can always find a line parallel ot x-axis, which intersects it at two points.

Hence, an injection graph is either an increasing or decreasing type. We can associate these characteristics with differential calculus. We can say that :

Either

$$\frac{dy}{dx} > 0 \quad \text{for all x}$$

or,

$$\frac{dy}{dx} < 0 \quad \text{for all x}$$

As a matter of fact the derivative can be equal to zero for certain values of "x" - not for an interval of "x". Thus, we can write the condition of increasing function : iif function is continuous and

$$\frac{dy}{dx} \geq 0 \quad \text{for all x;} \quad \text{equality holding for certain values of x}$$

Similarly, we can write the condition of decreasing function : iif function is continuous and

$$\frac{dy}{dx} \leq 0 \quad \text{for all x;} \quad \text{equality holding for certain values of x}$$

Many – one function

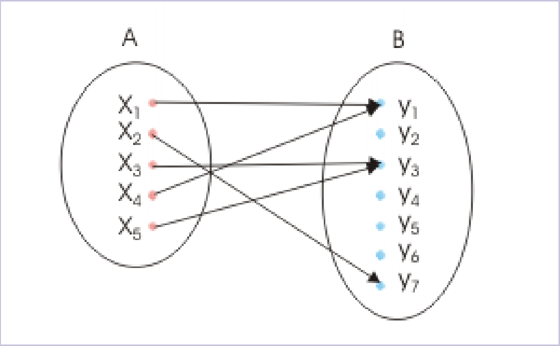
More than one pre-images of a function are related to same image.

Many - one function

> A function $f : A \rightarrow B$ is an many – one function, if two or more elements of domain set “A” have the same images in co-domain set “B”.

The test of condition for many-one function is easy : if a function is not one-one, then it is many-one. Alternatively, we can check literally going by the definition – whether there exist such many-one relation. A map diagram showing the relation will reveal such instances of many-one relation.

Many-one function



More than one pre-images of a function are related to same image.

Modulus function is one such many-one function. The function yields same value for positive and negative arguments of same magnitude.

$$f(x) = |x|$$

$$\Rightarrow f(-1) = |-1| = 1$$

$$\Rightarrow f(1) = |1| = 1$$

We should understand that a reverse function of the type “one to many” is barred from the very definition of function. The element of domain can be related to exactly one element in co-domain.

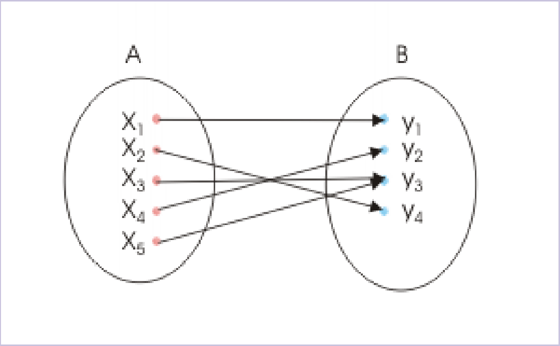
Onto function (surjection)

The definition of function puts the restriction on domain that every element in it is related. If we extend this restriction to co-domain also, then we get a function called “onto” or “surjection”.

Onto function (surjection)

> A function $f : A \rightarrow B$ is an onto function or surjection, if every element of co-domain set is the image of some element in the domain set “A”.

Onto function (surjection)



Every element of co-domain set is the image of some element in the domain set “A”.

One of the implications of surjection is that all elements of co-domain is related. It reduces the co-domain to range set. In other words, co-domain is also the range of the function.

$$\text{Co-domain of "f"} = \text{Range of "f"}$$

This equality of sets form one of the condition for testing a function to be surjection. Alternatively, we can check surjectivity by evaluating the rule of the function for the argument “x”. If the expression of “x” is valid for elements in co-domain, then the function is a surjection.

Example

Problem 2 : Consider a function defined as :

$$f : R \rightarrow R \quad by \quad f(x) = x^3 + 1 \quad \text{for all} \quad x \in R$$

Determine whether the function is a surjection?

Solution : We solve the rule for argument “x” as :

$$y = x^3 + 1$$

$$\Rightarrow x = (y - 1)^{1/3}$$

We see that expression on the right hand side is a valid real expression for all values of “y” in "R" i.e co-domain. Hence, given function is an onto function or surjection.

Into function

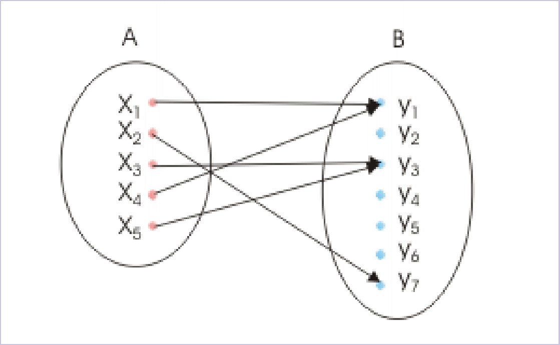
We have discussed in the beigining of this module that there is possibility that some of the elements of co-domains are not related. In that case, function is said to be into function.

Onto function (surjection)

> A function $f : A \rightarrow B$ is an into function, if there exists element in co-domain set, which has no pre-image in the domain set “A”.

One of the implications is that all elements of co-domain are not related to elements in domain set. In other words, range of the function is subset of the co-domain :

Into function



The range of the function is
subset of the co-domain .

$$\text{Range of "f"} \subset \text{Co-domain of "f"}$$

We can check whether a given function is an into function or not by checking whether the function is an onto set or not. If the function is not an onto function, then it an into function. Alternatively, we can check the equality of co-domain and range set. If they are not equal, then the function is into function.

Into function

Problem 3 : Consider a function defined as :

$$f : R \rightarrow R \quad by \quad f(x) = x^2 + 1 \quad forall \quad x \in R$$

Determine whether the function is an into function?

Solution : The rule of the function is :

$$y = x^2 + 1$$

The square of a real number is a positive number for all real number. Hence,

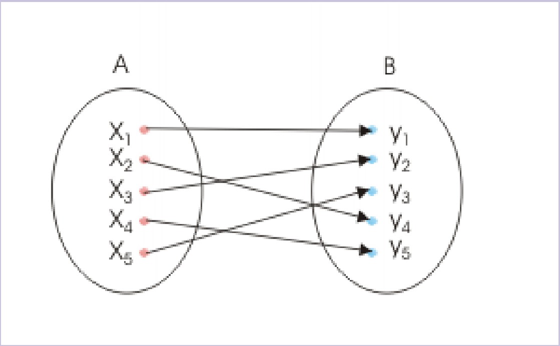
$$\Rightarrow y = x^2 + 1 > 0$$

It means that images are only the right half of the real number i.e. from zero to infinity. But, the co-domain of the function is “R”. It means that left half of the co-domain i.e. from negative infinity to less than zero has no image in “A”. Therefore, the given function is an into function.

One – one onto function (Bijection)

The bijection presents the most stringent condition for a function. Every element of both domain and co-domain is related to distinct element. This requirement is fulfilled when a function is both an injection and surjection.

One – one onto function (Bijection)



Every element of both domain and co-domain is related to distinct element.

The injection means that every element of domain is related to a distinct element in co-domain. On the other hand, surjection means that every element of co-domain is related, both distinctly and commonly. When conditions of injection and surjection are taken together, then it is also ensured that elements of co-domains are also related to distinct elements only.

One – one onto function (Bijection)

Problem 4 : Consider a function defined as :

$$f : A \rightarrow B \quad by \quad f(x) = \frac{x-2}{x-3}$$

Determine domain (A) and co-domain(B) of the function so that it is a bijection.

Solution : For determining domain of the function, we inspect the given rule,

$$f(x) = \frac{x-2}{x-3}$$

We observe that the given rational function is defined for all values of “x” except for x = 3. Hence, its domain is :

$$\text{Domain} = R - \{3\}$$

In order that the given function is a bijection, it should be both an injection and a surjection. For infectivity, we put :

$$f(x) = f(y)$$

$$\Rightarrow \frac{x-2}{x-3} = \frac{y-2}{y-3}$$

$$\Rightarrow (x-2)(y-3) = (x-3)(y-2)$$

$$\Rightarrow xy - 3x - 2y + 6 = xy - 2x - 3y + 6$$

$$\Rightarrow -3x - 2y = -2x - 3y$$

$$\Rightarrow x = y$$

Hence, function is an injection for the domain as determined above. Now, for surjection we solve the rule of the function for the argument (x),

$$\Rightarrow y = \frac{x-2}{x-3}$$

$$\Rightarrow xy - 3y = x - 2$$

$$\Rightarrow x(y-1) = 3y - 2$$

$$\Rightarrow x = \frac{3y-2}{y-1}$$

This equation is valid for all real values of “y” except “1”. Hence, function is surjection for all real values of “y” except for “1”. Hence, co-domain for the function to be a surjection is :

$$\text{Co-domain} = R - \{1\}$$

Thus, we conclude that the given function is bijection for the domain and co-domain as determined above.

$$\text{Domain} = R - \{3\}$$

$$\text{Co-domain} = R - \{1\}$$

Composition of functions

Function is a relation on two sets by a rule. It is a special mapping between two sets. It emerges that it is possible to combine two functions, provided co-domain of one function is domain of another function. The composite function is a relation by a new rule between sets, which are not common to the functions.

We can understand composition in terms of two functions. Let there be two functions defined as :

$$f : A \rightarrow B \quad \text{by} \quad \mathrm{f(x)} \text{ for all} \quad x \in A$$

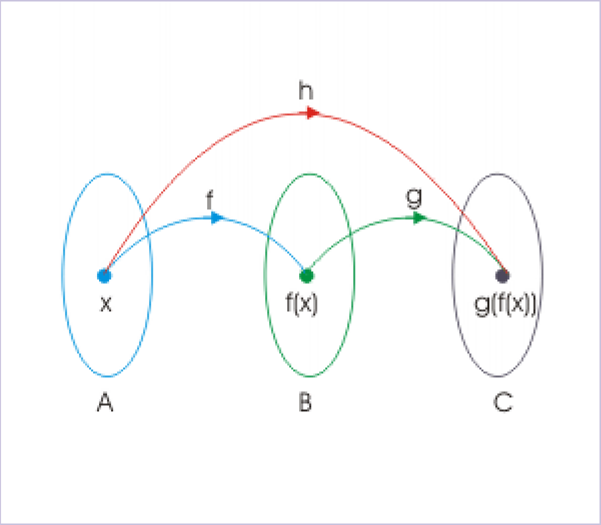
$$g : B \rightarrow C \quad \text{by} \quad \mathrm{g(x)} \text{ for all} \quad x \in B$$

Observe that set “B” is common to two functions. The rules of the functions are given by “f(x)” and “g(x)” respectively. Our objective here is to define a new function $h : A \rightarrow C$ and its rule.

Thinking in terms of relation, “A” and “B” are the domain and co-domain of the function “f”. It means that every element “x” of “A” has an image “f(x)” in “B”.

Similarly, thinking in terms of relation, “B” and “C” are the domain and co-domain of the function “g”. In this function, "f(x)" – which was the image of pre-image “x” in “A” – is now pre-image for the function “g”. There is a corresponding unique image in set "C". Following the symbolic notation, "f(x)" has image denoted by "g(f(x))" in "C". The figure here depicts the relationship among three sets via two functions (relations) and the combination function.

Composition of two functions



Composition functions is a special relation between sets not common to two functions.

For every element, “x” in “A”, there exists an element f(x) in set “B”. This is the requirement of function “f” by definition. For every element “f(x)” in “B”, there exists an element g(f(x)) in set “B”. This is the requirement of function “g” by definition. It follows, then, that for every element “x” in “A”, there exists an element g(f(x)) in set “C”. This concluding statement is definition of a new function :

$$h : A \to C \quad \text{by} \quad \mathrm{g(f(x))} \text{ for all} \quad x \in A$$

By convention, we call this new function as “gof” and is read as "g circle f" or "g composed with f".

$$gof(x) = g(f(x)) \quad \text{for all} \quad x \in A$$

The two symbolical representations are equivalent.

Example

Problem 1: Let two sets be defined as :

$$h : R \rightarrow R \quad \text{by} \quad x^2 \quad \text{for all} \quad x \in R$$

$$k : R \rightarrow R \quad \text{by} \quad x + 1 \quad \text{for all} \quad x \in R$$

Determine “hok” and “koh”.

Solution : According to definition,

$$hok(x) = h(k(x))$$

$$\Rightarrow hok(x) = h(x + 1)$$

$$\Rightarrow hok(x) = (x + 1)^2$$

Again, according to definition,

$$koh(x) = k(h(x))$$

$$\Rightarrow koh(x) = k\left(x^2\right)$$

$$\Rightarrow koh(x) = \left(x^2 + 1\right)$$

Importantly note that $hok(x) \neq koh(x)$. It indicates that composition of functions is not commutative.

Existence of composition set

In accordance with the definition of function, “f”, the range of “f” is a subset of its co-domain “B”. But, set “B” is the domain of function “g” such that there exists image g(f(x)) in “C” for every “x” in “A”. This means that range of “f” is subset of domain of “g” :

$$\text{Range of “f”} \subset \text{Domain of “g”}$$

Clearly, if this condition is met, then composition “gof” exists. Following this conclusion, “fog” will exist, if

$$\text{Range of “g”} \subset \text{Domain of “f”}$$

And, if both conditions are met simultaneously, then we can conclude that both “gof” and “fog” exist. Such possibility is generally met when all sets involved are set of real numbers, “R”.

Example

Problem 2: Let two functions be defined as :

$$f = \{(1,2), (2,3), (3,4), (4,5)\}$$

$$g = \{(2,4), (3,2), (4,3), (5,1)\}$$

Check whether “gof” and “fog” exit for the given functions?

Solution : Here,

$$\Rightarrow \text{Domain of “f”} = \{1,2,3,4\}$$

$$\Rightarrow \text{Range of “f”} = \{2,3,4,5\}$$

$$\Rightarrow \text{Domain of “g”} = \{2,3,4,5\}$$

$$\Rightarrow \text{Range of “g”} = \{4,2,3,1\} = \{1,2,3,4\}$$

Hence,

$$\Rightarrow \text{Range of “f”} \subset \text{Domain of “g”}$$

and

$$\Rightarrow \text{Range of “g”} \subset \text{Domain of “f”}$$

It means that both compositions “gof” and “fog” exist for the given sets.

Domain of Composition

Composition of two functions results in new rule for the new composite function. The expression of new rule may prohibit certain elements of the original domain set. For example, consider the function,

$$f(x) = \frac{1}{(1-x)} \quad \text{when} \quad x \neq 1$$

Clearly, the domain of function is R – {1}. Let us now see the expression of composition of function with itself,

$$\Rightarrow fof = f(f(x)) = \frac{1}{1 - \frac{1}{(1-x)}} = \frac{1-x}{1-x-1}$$

$$\Rightarrow fof = \frac{x-1}{x}$$

This expression is valid for real values of “x” when $x \neq 0$. Thus, we see that new rule has changed the domain of resulting function. The domain of the composition fof(x) is "R - {0,1}".

If functions “f” and “g” having different intervals of real numbers are involved in the composition, then we consider both the intervals and determine the domain of the composition by meeting requirement of both intervals (common interval). This aspect is illustrated in the examples given in the next section.

Algorithm for finding interval of composition

The most important aspect of working with composition of function is to combine intervals i.e. domains of two functions. Consider for example the function given by :

$$f(x) = x + 1; \quad 0 \le x \le 2$$

$$g(x) = |x|; \quad 0 \le x \le 3$$

Here, "|x|" is modulus function, which returns non-negative number for all real values of "x". We are required to find fog(x). What would be the domain of the resulting composition? Let us have a closer look at the definition of composition,

$$fog(x) = f(g(x)) = f(|x|)$$

In plain words, it means that argument of the function “f” is the function “g” itself. The function “f” is defined in the domain "0 ≤ x ≤ 2", whereas function “g” is defined in "0≤x≤3". From the expression of composition as above, it is clear that we need to ensure that value of “x” should lie in the domain interval of function “g”. Hence,

$$fog(x) = f(g(x)) = f(|x|); \quad 0 \le x \le 3$$

But function “f” is defined for values of “x”, which lie in its interval “0≤x≤2”. So when we expand the composition in accordance with the rule of function “f”, we should ensure that value of its argument – note that it is not the value of independent variable “x” - lies in the interval specified by its domain.

$$\Rightarrow fog(x) = |x| + 1; \quad 0 \le |x| \le 2 \quad \text{and} \quad 0 \le x \le 3$$

Note that interval of function “f” is written with respect to function “g” i.e. “|x|” – not “x”. Since requirements of both functions are required to be met simultaneously, the domain of the resulting composition is intersection of two domains. It is this reason that we use either “and” or a comma “,” to combine two intervals.

Now, we interpret the interval of modulus function “0≤ |x| ≤2”. We see that part of the interval, "|x| ≥0", is always true for all values of “x”. Whereas part of the interval, "|x| ≤2", means (we shall learn about interpreting modulus inequality in a separate module) :

$$-2 \le x \le 2$$

Combining intervals of two parts of the interval “0≤ |x| ≤2”, we conclude that it is equal to :

$$-2 \leq x \leq 2$$

Hence,

$$\Rightarrow fog(x) = |x| + 1; \quad -2 \leq x \leq 2 \quad \text{and} \quad 0 \leq x \leq 3$$

$$\Rightarrow fog(x) = |x| + 1; \quad 0 \leq x \leq 2$$

Examples

Problem 3: Consider the function as given here :

$$f(x) = x + 1; \quad 0 \leq x \leq 2$$

Determine fof(x).

Solution : The composition of the function with itself is :

$$fof(x) = f(f(x)) = f(x+1); \quad 0 \leq x \leq 2$$

$$\Rightarrow fof(x) = (x+1) + 1; \quad 0 \leq x + 1 \leq 2 \quad \text{and} \quad 0 \leq x \leq 2$$

$$fof(x) = x + 2; \quad -1 \leq x \leq 1 \quad \text{and} \quad 0 \leq x \leq 2$$

$$fof(x) = x + 2; \quad 0 \leq x \leq 1$$

Problem 4: A function is defined for real values by :

$$f(x) = \frac{1}{(1-x)} \quad \text{for all real values except} \quad x = 1$$

Determine f[f{f(x)}] and draw the graph of resulting composition.

Solution : This is composition triplet. We have already seen that :

$$\Rightarrow fof = f\left\{f(x)\right\} = \frac{x-1}{x} \quad \text{when} \quad x \neq 0, 1$$

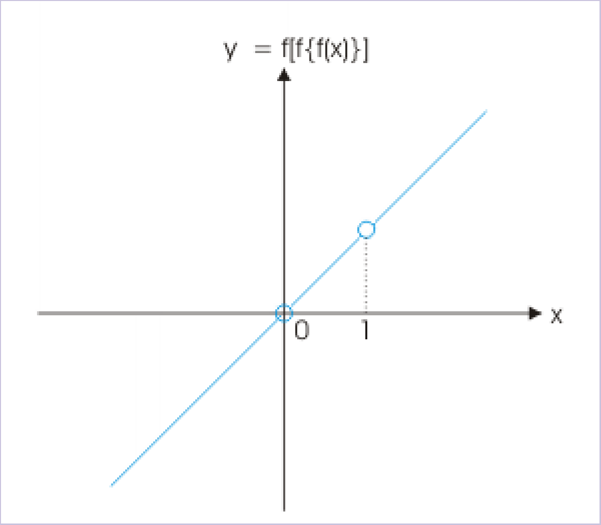
Compositing again with f(x), we have :

$$\Rightarrow f[f\{f(x)\}] = \frac{1}{1 - \frac{(x-1)}{x}} = \frac{x}{x - x + 1} \quad \text{when} \quad x \neq 0, 1$$

$$y = f[f\{f(x)\}] = x \quad \text{when} \quad x \neq 0, 1$$

The graph of the composition is as shown here :

Plot of compostion



The plot is a straight line with two undefined points.

Problem 5: A function is defined for real values by :

`| 1 + x ; 0 ≤ x ≤ 2 f(x) = | | 3 - x ; 2 < x ≤ 3`

Determine f(f(x)).

Solution : The function “f(x)” is combined with itself. Here,

`| f(1 + x) ; 0 ≤ x ≤ 2 f(f(x)) = | | f(3 - x) ; 2 < x ≤ 3`

We need to evaluate function for each of the above two intervals :

`f(1 + x) ; 0 ≤ x ≤ 2`

`| 1 + (1+x) ; 0 ≤ 1+ x ≤ 2 and 0 ≤ x ≤ 2 = | | 3 – (1+x); 2 < 1+ x ≤ 3 and 0 ≤ x ≤ 2`

`| 2 + x ; -1 ≤ x ≤ 1 and 0 ≤ x ≤ 2 = | | 2 – x ; 1 < x ≤ 2 and 0 ≤ x ≤ 2`

`| 2 + x ; 0 ≤ x ≤ 1 = | | 2 – x ; 1 < x ≤ 2`

Similarly,

`f(3 - x) ; 2 < x ≤ 3`

`| 1 + (3-x) ; 0 ≤ 3- x ≤ 2 and 2 < x ≤ 3 = | | 3 – (3-x) ; 2 < 3- x ≤3 and 2 < x ≤3`

`| 4 - x ; -3 ≤ x ≤ -1 and 2 < x ≤ 3 = | | x ; -1 < x ≤ 0 and 2 < x ≤ 3`

`| 4 - x ; 1 ≤ x ≤ 3 and 2 < x ≤ 3 = | | x ; 0 < x ≤ 1 and 2 < x ≤ 3`

`| 4 - x ; 2 < x ≤ 3 = | | x ; No common interval`

`= | 4 - x ; 2 < x ≤ 3`

Putting the results in the expression of “fof”, we have :

`| 2 + x ; 0 ≤ x ≤1 f(f(x)) = | 2 – x ; 1 < x ≤ 2 | 4 - x ; 2 < x ≤ 3`

Properties of composition

The composition is generally not commutative except for some special functions.

$$gof(x) \neq fog(x)$$

On the other hand, composition among three functions is independent of parentheses and hence is associative.

$$(gof)oh = go(foh)$$

Inverse of a function

Inverse relation is like looking a relation in opposite direction. Equivalently, it is also like an image in a mirror. For example, consider the relation “husband of”. The inverse to this relation is “wife of”. This is an explicit relation very easily conceivable. In other situations involving function, inverse relations may not be so explicit. We shall, therefore, develop mathematical technique to obtain inverse function (relation) for a given function (relation).

In order to facilitate easy recapitulation of concepts and terms for the study of inverse relation, we can refer meaning attached to following terms :

- f : denotes function relation from domain “A” to co-domain “B” of “f” (implicit inference of sets).
- $f: A \rightarrow B$: denotes function relation from domain “A” to co-domain “B” of “f” (explicit reference of sets).
- $f(x)$: denotes image, an instance of function “f”, image under “f”, rule of “f”.
- f^{-1} : denotes inverse function relation from domain “A” to co-domain “B” of “ f^{-1} ” (implicit reference of sets)
- $f^{-1}: A \rightarrow B$: denotes inverse function relation from domain “A” to co-domain “B” of “ f^{-1}” (explicit reference of sets)
- $f^{-1}(x)$: denotes pre-image, an instance of function “ f^{-1}”, image under “ f^{-1}”, rule of “ f^{-1}”

Inverse of an element

We use the concept of pre-image and image to connect the elements of a function in the direction from domain “A” to co-domain “B”. The related elements are connected by a rule “f(x)” such that :

$$\text{Image} = f(x) = f(\text{pre-image})$$

Clearly, “x” is the pre-image and “f(x)” is image. Now, we want to derive a similar rule, " $f^{-1}(x)$ ", which evaluates to pre-image like :

$$\text{Pre-image(s)} = f^{-1}(x) = f^{-1}(\text{image})$$

Clearly, “x” is now the image and “ $f^{-1}(x)$ ” is pre-image(s). The important point to understand here is that the image in the co-domain set can be related to none, one or more elements in domain set. Therefore, this rule may evaluate accordingly to value(s) – none, one or more - for the pre-images.

Method to construct an inverse rule

We construct an inverse rule in step-wise manner as enumerated here with an example :

Step 1: Write down the rule of the given function “f”.

Let the given rule be f(x) given by :

$$f(x) = x^2$$

Let us put y = f(x). Then,

$$\Rightarrow y = f(x) = x^2$$

This relation gives us one value of image. For example, if x = 3, then

$$\Rightarrow y = 3^2 = 9$$

Step 2: Solve for “x”

$$\Rightarrow x = \pm\sqrt{y}$$

Step 3: Replace “x” which represents pre-image by the symbol “ $f^{-1}(x)$” and replace “y” which represents image by “x”. For the given function, $f\left(x\right) = x^2$, the new inverse rule is :

$$\Rightarrow f^{-1}(x) = \pm\sqrt{x}$$

This is how we construct the inverse rule. Note emphatically that “x” now represents “image” and “ $f^{-1}(x)$ ” represents “pre-image”. For example, if image is “9”, then we can find its pre-image (s), using this new rule as :

$$\Rightarrow f^{-1}(9) = \pm\sqrt{9} = \pm 3$$

Thus, the required pre-images is a set of two pre-images :

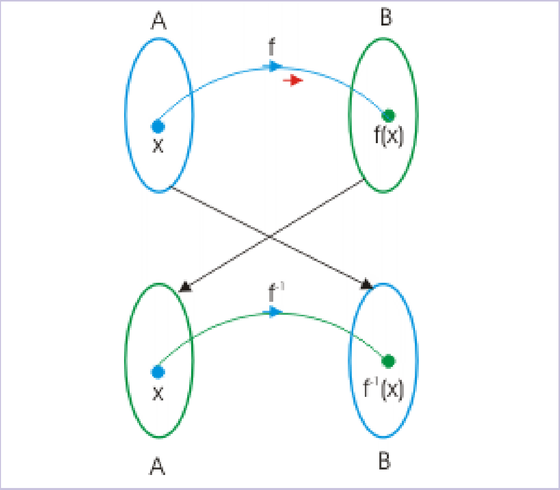
$$\Rightarrow f^{-1}(9) = \{-3,3\}$$

Inverse of a function

Once the inverse rule is constructed, it is easy to define inverse function. However, we should be careful in one important aspect. An inverse function, " f^{-1} " is a function ultimately. This puts the requirement that every element of the domain of the new function " f^{-1} " should be related to exactly one element to its co-domain set.

We must also understand that this new function, " f^{-1} ", gives the perspective of relation from co-domain to domain of the given function "f". However, new function " f^{-1} " is read from its new domain to its new co-domain. After all this is how a function is read. This simply means that domain and co-domain of the function "f" is exchanged for " f^{-1} ".

Inverse function



The domain and co-domain sets are exchanged.

Further, inverse function is inverse of a given function. Again by definition, every element of domain set of the given function "f" is also related to exactly one element of in its co-domain. Thus, there is bidirectional requirement that elements of one set are related to exactly one element of other set. Clearly, this requirement needs to be fulfilled, before we can define inverse function.

In other words, we can define inverse function, " f^{-1} ", only if the given function is an injection and surjection function (map or relation) at the same time. Hence, iff function, “f” is a bijection, then inverse function is defined as :

$$f^{-1}: A \rightarrow B \quad \text{by} \quad f^{-1}(x) \quad \text{for all} \quad x \in A$$

We should again emphasize here that sets “A” and “B” are the domain and co-domain respectively of the inverse function. These sets have exchanged their place with respect to function “f”. This aspect can be easily understood with an illustration. Let a function “f” , which is a bijection, be defined as :

Let A = {1,2,3,4} and B={3,6,9,12}

$$f: A \rightarrow B \quad \text{by} \quad f(x) = 3x \quad \text{for all} \quad x \in A$$

Then, the function set in the roaster form is :

$$\Rightarrow f = \{(1,3),(2,6),(3,9),(4,12)\}$$

This function is clearly a bijection as only distinct elements of two sets are paired. Its domain and co-domains are :

$$\Rightarrow \text{Domain of “f”} = \{1,2,3,4\}$$

$$\Rightarrow \text{Co-domain of “f”} = \{3,6,9,12\}$$

Now, the inverse function is given by :

$$f^{-1}: A \rightarrow B \quad by \quad f(x) = \frac{x}{3} \quad \text{for all} \quad x \in A$$

$$\text{where} \quad A = \{3,6,9,12\}$$

In the roaster form, the inverse function is :

$$\Rightarrow f^{-1} = \left\{(3,1),(6,2),(9,3),(12,4)\right\}$$

Note that we can find inverse relation by merely exchanging positions of elements in the ordered pairs. The domain and co-domain of new function “ f^{-1} ” are :

$$\Rightarrow \text{Domain of} \quad f^{-1} = \left\{3,6,9,12\right\}$$

$$\Rightarrow \text{Co}-\text{domain of} \quad f^{-1} = \left\{1,2,3,4\right\}$$

Thus, we see that the domain of inverse function " f^{-1}" is co-domain of the function "f" and co-domain of inverse function " f^{-1}" is domain of the function "f".

Example

Problem 1: A function is given as :

$$f : R \rightarrow R \quad \text{by} \quad f(x) = 2x + 5 \quad \text{for all} \quad x \in R$$

Construct the inverse rule. Determine f(x) for first 5 natural numbers. Check validity of inverse rule with the values of images so obtained. Find inverse function, if it exists.

Solution : Following the illustration given earlier, we derive inverse rule as :

$$y = 2x + 5$$

$$\Rightarrow x = \frac{y-5}{2}$$

Changing notations,

$$\Rightarrow f^{-1}(x) = \frac{x-5}{2}$$

The images i.e. corresponding f(x), for first five natural numbers are :

$$f(1) = 2x + 5 = 7; \quad f(2) = 9; \quad f(3) = 11; \quad f(4) = 13 \quad and \quad f(5) = 15$$

Now, the corresponding pre-images, using inverse rule for two values of images are :

$$\Rightarrow f^{-1}(7) = \frac{7-5}{2} = 1$$

$$\Rightarrow f^{-1}(11) = \frac{11-5}{2} = 3$$

Thus, we see that the inverse rule correctly determines the pre-images as intended. Now, in order to find inverse function, we need to determine that the given function is an injection and surjection. For injection, let us assume that “x_1 ” and “x_2” be two different elements such that :

$$f(x_1) = f(x_2)$$

$$\Rightarrow 2x_1 + 5 = 2x_2 + 5$$

$$x_1 = x_2$$

This means that given function is an injection. Now, to prove surjection, we solve the rule for “x” as :

$$\Rightarrow x = \frac{y-5}{2}$$

We see that this equation is valid for all values of “R” i.e. all values in the co-domain of the given function. This means that every element of the co-domain is related. Hence, given function is surjection. The inverse function, therefore, is given as :

$$f^{-1} : R \rightarrow R \quad by \quad f(x) = \frac{(x-5)}{2} \quad \text{for all} \quad x \in R$$

Properties of inverse function

There are few characteristics of inverse function that results from the fact that it is inverse of a bijection. We can check the validity of these properties in terms of the example given earlier. Let us define a bijection function as defined earlier :

$$\text{Let} \quad A = \{1,2,3,4\} \quad \text{and} \quad B = \{3,6,9,12\}$$

$$f : A \rightarrow B \quad \text{by} \quad f(x) = 3x \quad \text{for all} \quad x \in A$$

Inverse function is unique function

This means that there is only one inverse function. For the given function the inverse function is :

$$f^{-1}: A \rightarrow B \quad by \quad f(x) = \frac{x}{3} \quad forall \quad x \in A$$

$$\text{where} \quad A = \{3,6,9,12\}$$

In the roaster form, the inverse function is :

$$\Rightarrow f^{-1} = \left\{(3,1), (6,2), (9,3), (12,4)\right\}$$

This inverse function is unique to a given bijection.

Inverse function is bijection

We see that inverse comprises of ordered pairs such that elements of domain and co-domain are distinctly related to each other.

$$\Rightarrow f^{-1} = \left\{(3,1), (6,2), (9,3), (12,4)\right\}$$

This mean that the inverse function is bijection.

Graph of inverse function

If a function is bijection, then the inverse of function exists. On the other hand, a function is bijection, if it is both one-one and onto function. We know that one-one function is strictly monotonic in its domain. Hence, an onto function is invertible, if its graph is strictly monotonic i.e. either increasing or decreasing.

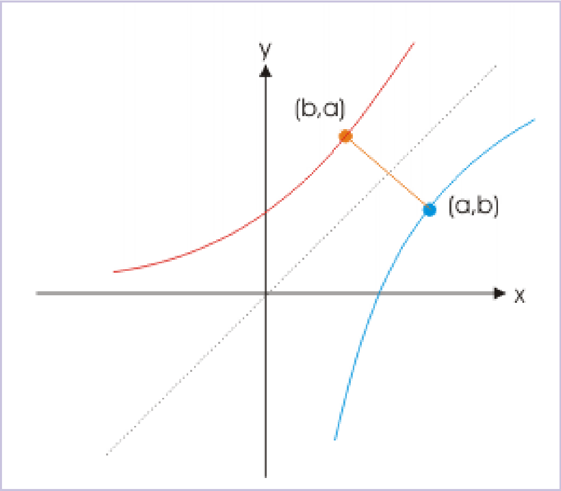
In order to investigate the nature of the inverse graph, let us consider a plot of an invertible function, “f(x)”. Let (a,b) be a point on the plot. Then, by definition of an inverse function, the point (b,a) is a point on the plot of inverse function, if plotted on the same coordinate system.

$$y = f(x)$$

$$y = f^{-1}\left(x\right)$$

By geometry, the line joining points (a,b) and (b,a) is bisected at right angles by the line y = x. It means that two points under consideration are object and image for the mirror defined by y=x. This relationship also restrains that two plots can intersect only at line y = x.

Graph of inverse function



The points on two plots are object and image for the mirror defined by y=x.

Example

Problem 2 : Two functions, inverse of each other, are given as :

$$f(x) = x^2 - x + 1$$

$$f^{-1}(x) = \frac{1}{2} + \sqrt{\left(x - \frac{3}{4}\right)}$$

Find the solution of the equation :

$$x^2 - x + 1 = \frac{1}{2} + \sqrt{\left(x - \frac{3}{4}\right)}$$

Solution :

Statement of the problem : The given functions are inverse to each other, which can intersect only at line defined by y = x. Clearly, the intersection point is the solution of the equation.

$$y = f(x) = x$$

$$\Rightarrow x^2 - x + 1 = x \quad \Rightarrow x^2 - 2x + 1 = 0$$

$$\Rightarrow (x-1)^2 = 0$$

$$\Rightarrow x = 1$$

This is the answer. It is interesting to know that we can also proceed to find the solution by working on the inverse function. This should also give the same result as given functions are inverse to each other.

$$y = f^{-1}(x) = x$$

$$\Rightarrow \frac{1}{2} + \sqrt{\left(x - \frac{3}{4}\right)} = x$$

$$\Rightarrow \sqrt{\left(x - \frac{3}{4}\right)} = x - \frac{1}{2}$$

Squaring both sides,

$$\Rightarrow \left(x - \frac{3}{4}\right) = \left(x - \frac{1}{2}\right)^2 = x^2 + \frac{1}{4} - x$$

$$\Rightarrow x^2 + \frac{1}{4} - x - x + \frac{3}{4} = 0 \quad \Rightarrow x^2 - 2x + 1 = 0$$

$$\Rightarrow x = 1$$

Domain and range

We have seen that a function is a special relation. In the same sense, real function is a special function. The special about real function is that its domain and range are subsets of real numbers “R”. In mathematics, we deal with functions all the time – but with a difference. We drop the formal notation, which involves its name, specifications of domain and co-domain, direction of relation etc. Rather, we work with the rule alone. For example,

$$f(x) = x^2 + 2x + 3$$

This simplification is based on the fact that domain, co-domain and range are subsets of real numbers. In case, these sets have some specific intervals other than “R” itself, then we mention the same with a semicolon (;) or a comma(,) or with a combination of them :

$$f(x) = \sqrt{(x+1)^2 - 1}; x < -2, x \geq 0$$

Note that the interval “ $x < -2, x \geq 0$ ” specifies a subset of real number and defines the domain of function. In general, co-domain of real function is “R”. In some cases, we specify domain, which involves exclusion of certain value(s), like :

$$f(x) = \frac{1}{1-x}, x \neq 1$$

This means that domain of the function is $R - \{1\}$. Further, we use a variety of ways to denote a subset of real numbers for domain and range. Some of the examples are :

- $x > 1$: denotes subset of real number greater than “1”.
- $R - \{0, 1\}$ denotes subset of real number that excludes integers “0” and “1”.
- $1 < x < 2$: denotes subset of real number between “1” and “2” excluding end points.
- $(1{,}2]$: denotes subset of real number between “1” and “2” excluding end point “1”, but including end point “2”.

Further, we may emphasize the meaning of following inequalities of real numbers as the same will be used frequently for denoting important segment of real number line :

- Positive number means x > 0 (excludes “0”).
- Negative number means x < 0 (excludes “0”).
- Non - negative number means x ≥ 0 (includes “0”).
- Non – positive number means x ≤ 0 (includes “0”).

Domain of real function

Domain of real function is a subset of “R” such that rule “f(x)” is real. This concept is simple. We need to critically examine the given function and evaluate interval of “x” for which “f(x)” is real.

In this module, we shall restrict ourselves to algebraic functions. We determine domain of algebraic function for being real in the light of following facts :

- If the function has rational form p(x)/q(x), then denominator q(x) ≠ 0.
- The term √x is a positive number, where x>=0.
- The expression under even root should be non-negative. For example the function $\sqrt{(x^2+3x-2)}$ to be real, $x^2+3x-2 \geq 0$.
- The expression under even root in the denominator of a function should be positive number. For example the function $\frac{1}{\sqrt{(x^2+3x-2)}}$ to be real, $x^2+3x-2>0$. Note that zero value of expression is not permitted in the denominator.

Here, we shall work with few examples as illustration for determining domain of real function.

Examples

Problem 1 : A function is given by :

$$f(x) = \frac{1}{x+1}$$

Determine its domain set.

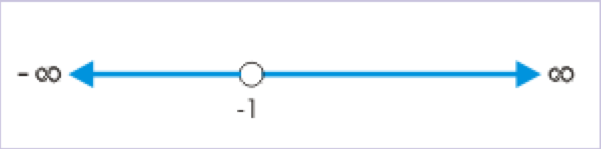
Solution : The function, in the form of rational expression, needs to be checked for its denominator. The denominator should not evaluate to zero as “a/0” form is undefined. For given function in the question, the denominator evaluates to zero when,

$$x + 1 = 0$$

$$\Rightarrow x = -1$$

Hence, domain of the given function is $R - \{-1\}$. The representation of the domain on real number line is shown with a dark line on either side of the excluded point “-1”.

Domain of a function



The number "-1" is excluded from the domain.

Problem 2 : A function is given by :

$$f(x) = \sqrt{(x^2 - 5x + 6)}$$

Determine its domain set.

Solution : We observe that given function is a square root of a quadratic polynomial. The expression within square root should be a non-negative number as square root of a negative number is not a real number. It means that expression under even root should be non-negative.

We factorize the quadratic expression in order to find corresponding interval for which expression under the root is non-negative.

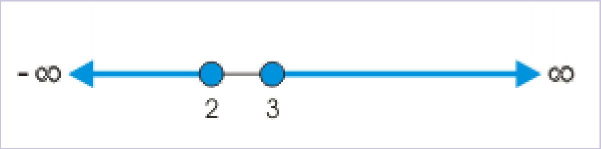
$$\Rightarrow \left(x^2 - 5x + 6\right) \geq 0$$

$$\Rightarrow (x-2)(x-3) \geq 0$$

Note: There are two specific sign schemes. These schemes are very helpful in determining interval for inequalities. We shall discuss and use them in next module. For the present, we carry on with general interpretation so that we realize difficulties in estimating intervals of inequalities without sign schemes and also that we have insight into the requirements of interval formation.

We interpret this result in reference to quadratic equation. When x ≤ 2, both factors of quadratic equation are non-positive and their product is non-negative. We can check with one such value like “1” or “-1”. On the other hand, when x ≥ 3, both factors are non-negative and their product is also non-negative. It turns out that these two conditions correspond to two intervals, which are disconnected to each other.

Domain of a function



The domain consists of two disjointed intervals.

The representation of the domain interval on real number line is shown with thick line and small filled circles. From the representation on the figure also, it is clear that it is a case of two disjoint intervals. We, therefore, represent the valid domain of the function with the help of the concept of union of two sets (intervals) in the following manner :

$$-\infty < x \le 2 \quad \text{or} \quad 3 \le x < \infty$$

$$\Rightarrow -\infty < x \le 2 \quad \cup \quad 3 \le x < \infty$$

i.e.

$$\Rightarrow (-\infty, 2] \quad \cup \quad [3,\infty)$$

This interpretation is typical of product of two linear factors, which is greater than or equal to zero. This interpretation, as a matter of fact, can be used as an axiom in general for deciding interval, involving product of two factors.

Range of a real function

Range is a set of images. It is a subset of co-domain. The requirement, here, is to find the interval of the co-domain for which there is “pre-image” in the domain set. In other words, we need to find the values of “y” within the domain of the function.

Further, we have already developed technique to find the inverse element i.e. pre-images, while studying inverse function. We shall apply the same concept here to decide range of the function. However, unlike determining domain, it is extremely helpful that we follow a step-wise algorithm to determine the range. It is given here as :

1. Find domain.
2. Put y = f(x).
3. Solve the function for “x” in terms of “y”.
4. Find the values of “y” for which “x” is real in the domain of the function determined in step 1.

While determining range of a function, we need to be careful with regard to two important aspects :

1: The values obtained for range are consistent with the function. This means that we should check the range against the requirement of a given function. For example, if range of a square root function is evaluated as say [-3,3], then we need to discard negative interval. A square root can not be negative. Hence, the correct range would be [0,3].

2: We need to exclude values of function (y) corresponding to invalid values of “x”. This is particularly the case if domain is a continuous interval except few points barred by the definition of the function.

The best way to understand this algorithm is to work with few examples.

Examples

Problem 3 : A function is given by :

$$y = \sqrt{(9 - x^2)}$$

Determine its range.

Solution : For real value of “y”, the expression $\left(9 - x^2\right)$ is non-negative number. It means that :

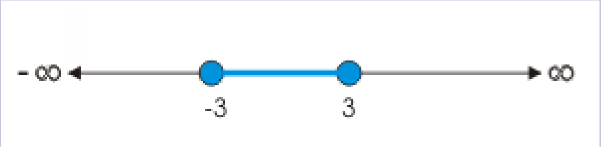
$$\Rightarrow 9 - x^2 \ge 0$$

$$\Rightarrow x^2 - 9 \le 0$$

$$\Rightarrow (x-3)(x+3) \le 0$$

We interpret this result in reference to the given quadratic equation. When $x \ge -3, \quad \text{but} \quad x \le 3$, the signs of two factors are opposite and hence their product is less than or equal to zero. Outside this interval, the expression evaluates to positive number.

Domain of a function



The domain lies between two end points, inclusive of them.

The representation of the domain interval on real number line is shown with thick line and two small filled circles. We see that real values of “x” lies between “-3” and “3”, including end points. We represent the valid domain as :

$$-3 \leq x \leq 3$$

or

$$[-3,3]$$

This interpretation is typical of product of two linear factors, which is less than or equal to zero. This interpretation can also be used as an axiom in general for deciding interval, involving two factors.

In order to find range, we solve the function for “x”,

$$y = \sqrt{(9 - x^2)}$$

$$\Rightarrow x = \pm\sqrt{(9 - y^2)}$$

Following the same analysis as for domain, we reach the conclusion that the value of “y” for real “x” is given by the interval :

$$[-3,3]$$

Now, the examination of given function reveals that “y” can be only positive number (note that no negative sign precede square root expression) in the expression for "Y" :

$$y = \sqrt{(9 - x^2)}$$

Hence, “y” can not be negative. Note that we determined interval of "y", which includes negative numbers also. Thus, we conclude that range of the given function is half of the interval obtained earlier, which includes zero also :

$$[0,3]$$

Problem 4 : A function is given by :

$$y = \frac{x^2}{1 + x^2}$$

Determine its range.

Solution : We observe that the both numerator and denominator of the given function are non-negative. It is because “ x^2 “ always evaluates to non-negative number. It means that given function is real for all values of “x”. Thus, domain of function is “R”.

In order to find range, we solve the function for “x”,

$$y = \frac{x^2}{1 + x^2}$$

$$\Rightarrow yx^2 + y = x^2$$

$$\Rightarrow x^2 = \frac{y}{1 - y}$$

$$\Rightarrow x = \pm\sqrt{\left(\frac{y}{1-y}\right)}$$

For “x” to be real, the expression within square root should be non - negative. This case, however, is different in that it is a ratio of two linear expressions. It is possible that denominator is positive, but numerator is negative or vice - versa. As such, the rational expression as a whole will be negative. In the nutshell, we need to evaluate “x” for following requirements (Note : we are presenting basic reasoning here. Subsequently, we will learn more sophisticated means to determine valid intervals of variables) :

1. Total expression within the square root as a whole is non-negative number as square root of a negative number is not a real number.
2. For positive value of "y" in the numerator, the denominator is non-negative as square root of a negative number is not a real number.
3. The denominator does not evaluate to zero. The form “y/0” is undefined.

For the first requirement, the expression in the square root should be greater than or equal to zero i.e non-negative number.

$$\frac{y}{1-y} \geq 0$$

$$\Rightarrow y \geq 0$$

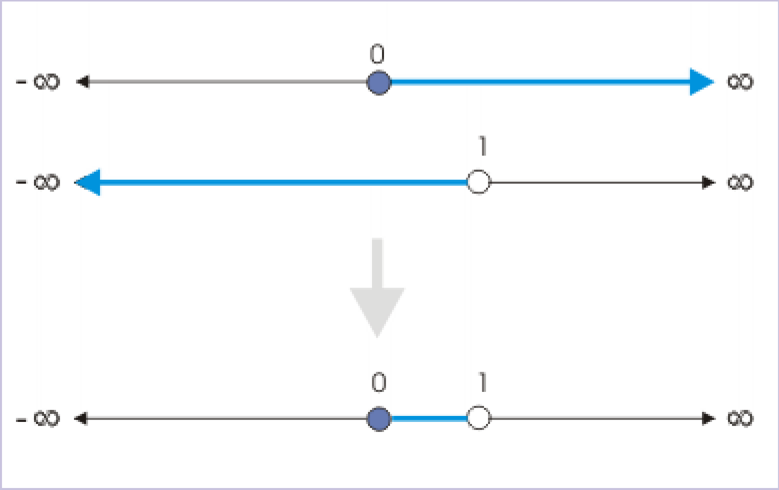
Further, denominator of the expression “1-y” is non-negative. Also, “1-y” is not equal to zero. Combining two requirements, the expression is a positive number :

$$1 - y > 0$$

$$\Rightarrow y < 1$$

Combining two intervals i.e. intersection of two intervals, we have the range of the function as :

Range of a function



Range of the function is equal to intersection of two intervals.

$$\Rightarrow 0 \leq y < 1$$

Classification of functions

In mathematics, we deal with specific real functions, which are characterized by specific domain, range and rules. Some of the familiar functions are polynomial, rational, irrational, trigonometric, exponential, logarithmic functions and piece wise defined functions etc. These functions are further combined to form more complex function following certain definition or rule so that function is meaningful for real values.

There are varieties of functions. These functions are broadly classified under three headings :

- **Algebraic functions :** polynomial, rational and irrational functions
- **Transcendental functions :** trigonometric, inverse trigonometric, exponential and logarithmic functions
- **Piece wise defined functions :** modulus, greatest integer, least integer, fraction part functions and other specific piece wise definitions

Polynomial function is further classified based on (i) numbers of terms eg. monomial, binomial, trinomial etc. (ii) numbers of variables involved eg. function in one or two variables and (iii) degree of the polynomial eg. linear, quadratic, cubic, bi-quadratic etc.

Generally we deal with function expressions in one variable in which dependent variable (y) is explicitly related to independent variable (x). Such functions are called explicit function.

$$y = x^2 - 2x + 1$$

On the other hand, there are function rules in which “y” is not explicitly related to “x”. Such functions are called implicit functions.

$$\sin\left(x^2 + xy + y^2\right) = xy$$

Further, we also use properties of function like periodicity (repetition of function values at regular intervals of independent variable) and polarity (odd or even) to characterize a function. For this reason, we sometimes name a function like periodic, non-periodic, aperiodic, odd, even or equal function etc.

Important concepts

In this section, we discuss few important concepts, which are frequently used in determining domain and range.

Defined and undefined expressions

Defined expressions are meaningful and unambiguous. On the contrary, undefined expressions are not meaningful. Most of the undefined expressions results from combination of zeros and infinity in various ways. There is, however, no unanimity about “undefined” values among mathematicians. Hence, we shall present two lists – one list which is undefined in all contexts and another list which may be defined in certain context. We consider this later list as defined expressions, unless otherwise stated.

Undefined in all contexts

$$\frac{x}{0}, \quad \infty - \infty, \quad (-1)^{\infty}, \quad \frac{\infty}{\infty}, \quad 0X\left(-\infty\right)$$

We should understand here that an expression is undefined when it can not be interpreted. The important point is that it has got nothing to do with the magnitude of quantity. Emphatically, infinity is not undefined. We shall discuss this aspect subsequently.

Note that " $(-1)^{\infty}$ " is undefined, because it is not certain whether the expression will evaluate to "-1" or "1". On the other hand, expression " $(1)^{\infty}$ " is defined because we are sure that the expression will evaluate to "1", however large is infinity.

Defined in some contexts

$$0^0 = \text{undefined or } 1$$

$$\infty^0 = \text{undefined or } 1$$

$$1^{\infty} = \text{undefined or } 1$$

$$0X\infty = \text{undefined or } 0$$

For our purpose, unless otherwise stated, we shall consider later set as defined.

Infinity

Infinity is not a member of real number set “R”. Strictly we can not write infinity like :

$$x = \infty, \quad \text{where “x” is real.}$$

For this reason, the interval of real number set is defined in terms of infinity without equality as :

$$-\infty < x < \infty \quad \text{or} \quad (-\infty, \infty)$$

We may emphasize that infinity by itself is “unbounded” – not undefined. What it means that we can interpret infinity – even though its value is not known. We can say it is very large number (either positive or negative as the case be), but we can not interpret undefined values at all.

It is easy to interpret operations with infinity. We need to only keep the meaning of infinity as a very large number in focus. Various operations, involving infinity are presented here :

1: The plus or minus infinity is not changed by adding or subtracting real number.

$$\infty \pm x = \infty$$

$$-\infty \pm x = -\infty$$

Above results are on expected line. Addition or subtraction of finite values will only yield a large number. It is so because infinity can be greater than a large value that we might conceive.

2: Addition of two infinities is infinity.

$$\infty + \infty = \infty$$

3: Difference of two infinities is undefined.

$$\infty - \infty = \text{Undefined}$$

Addition of two infinities is definitely a very large number. We are, however, not sure about their difference. The difference of two infinities can either be small or large. It depends on the relative "largeness" of two infinities. Hence, difference of two infinities is "undefined".

4: Product of two infinities are inferred as :

$$\infty X \infty = \infty$$

$$-\infty X \infty = -\infty$$

$$-\infty X - \infty = \infty$$

5: Product of infinity with a real number “x” is given as :

$$x X \infty = \infty, \quad \text{if} \quad x > 0$$

$$x X \infty = -\infty, \quad \text{if} \quad x < 0$$

$$x X \infty = 0, \quad \text{if} \quad x = 0$$

6: Division of infinity by infinity is not defined.

$$\frac{\infty}{\infty} = \text{Undefined}$$

7: A real number, “x”, raised to infinity

$$x^{\infty} = 0, \quad \text{if} \quad 0 < x < 1$$

$$x^{\infty} = 0, \quad \text{if} \quad x = 0$$

$$x^{\infty} = \infty, \quad \text{if} \quad x > 1$$

8: An infinity raised to infinity is defined.

$$\infty^{\infty} = \infty$$

Inequality

Inequality is an important concept in understanding function and its properties – particularly domain and range. Many function forms are valid in certain interval(s) of real numbers. This means definition of function is subjected to certain restriction of values with respect to dependent and independent variables. The restriction is generally evaluated in terms of algebraic inequalities, which may involve linear, quadratic, higher degree polynomials or rational polynomials.

Function definition and inequality

A function imposes certain limitations by virtue of definition itself. We have seen such restriction with respect to radical functions in which polynomial inside square root needs to be non-negative. We have also seen that denominator of a rational function should not be zero. We shall learn about different functions in subsequent modules. Here, we consider few examples for illustration :

1 : $f(x) = \log_a\left(3x^2 - x + 4\right)$

Here, logrithmic function is defined for $a \in (0,1) \cup (1,\infty)$ and

$$3x^2 - x + 4 > 0$$

2 : $f(x) = \arcsin\left(3x^2 - x + 4\right)$

Here, arcsine function is defined in the domain [-1,1]. Hence,

$$-1 \le 3x^2 - x + 4 \le 1$$

It is clear that we need to have clear understanding of algebraic inequalities as function definitions are defined with certain condition(s).

Forms of function inequality

Function inequality compares function to zero. There are four forms :

1 : f(x) < 0

2 : f(x) ≤ 0

3 : f(x) > 0

4 : f(x) ≥ 0

Here, f(x) < 0 and f(x) >0 are strict inequalities as they confirm the notion of “less than” and “greater than”. There is no possibility of equality. If a strict inequality is true, then non-strict equality is also true i.e.

1 : If f(x) > 0 then f(x) ≥ 0 is true.

2 : If f(x) ≥ 0 then f(x) > 0 is not true.

3 : If f(x) < 0 then f(x) ≤ 0 is true.

4 : If f(x) ≤ 0 then f(x) < 0 is not true.

Further, we may be presented with inequality which compares function to non-zero value :

$$3x^2 - x \leq -4$$

However, such alterations are equivalent expressions. We can always change this to standard form which compares function with zero :

$$3x^2 - x + 4 > 0$$

Inequalities

Some important definitions/ results are enumerated here :

- Inequalities involve a relation between two real numbers or algebraic expressions.
- The inequality relations are "<", ">", "≤" and "≥".
- Equal numbers can be added or subtracted to both sides of an inequality.

- Both sides of an inequality can be multiplied or divided by a positive number without any change in the inequality relation.
- Both sides of an inequality can be multiplied or divided by a negative number with reversal of inequality relation.
- Both sides of an inequality can be squared, provided expressions are non-negative. As a matter of fact, this conclusion results from rule that we can multiply both sides with a positive number.
- When both sides are replaced by their inverse, the inequality is reversed .

Equivalently, we may state above deductions symbolically.

$$\text{If} \quad x > y, \quad \text{then :}$$

$$x + a > y + a$$

$$ax > ay; \quad a > 0$$

$$ax < ay; \quad a < 0$$

$$x^2 > y^2; x, y > 0$$

$$\frac{1}{x} < \frac{1}{y}; \quad \text{when “x” and “y” have same sign.}$$

It is evident that we can deduce similar conclusions with the remaining three inequality signs.

Intervals with inequalities

In general, a continuous interval is denoted with "less than (<)" or "less than equal to (≤)" inequalities like :

$$1 < x \le 5$$

The segment of a real number line from a particular number extending to plus infinity is denoted with “greater than” or “greater than equal to” inequalities like :

$$x \geq 3$$

The segment of real number line from minus infinity to a certain number on real number line is denoted with “less than(<) or less than equal to (≤)” inequalities like :

$$x \leq -3$$

Two disjointed intervals are combined with “union” operator like :

$$1 < x \leq 2 \quad \cup \quad x > 5$$

Linear inequality

Linear function is a polynomial of degree 1. A linear inequality can be solved for intervals of valid “x” and “y” values, applying properties of inequality of addition, subtraction, multiplication and division. For illustration, we consider a logarithmic function, whose argument is a linear function in x.

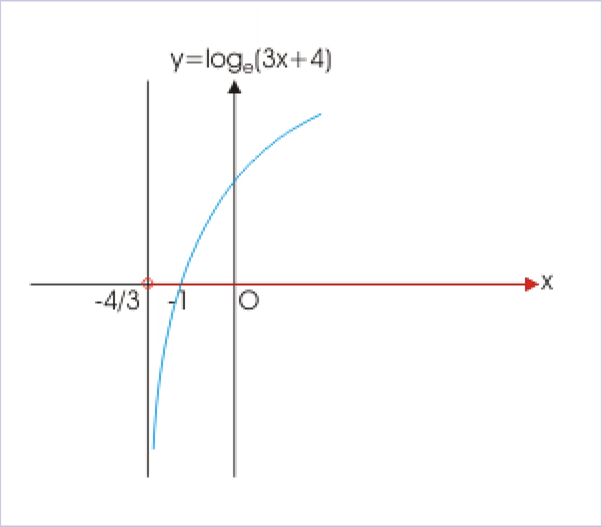
$$f(x) = \log_e(3x + 4)$$

The argument of logarithmic function is a positive number. Hence,

$$\Rightarrow x > -\frac{4}{3}$$

Therefore, interval of x i.e. domain of logarithmic function is $(-4/3, \infty)$. The figure shows the values of “x” on a real number line as superimposed on x-axis. Note x= - 4/3 is excluded.

Graph of logarithmic function



Domain is traced on x-axis.

When f(x) = 0,

$$3x+4=e^{f(x)}=e^{0}=1$$

$$\Rightarrow x=-1$$

It means graph intersects x-axis at x=-1 as shown in the figure. From the figure, it is clear that range of function is real number set R.

Note: We shall similarly consider inequalities involving polynomials of higher degree, rational function etc in separate modules.

Example:
Problem : A linear function is defined as f(x)=2x+2. Find valid intervals of "x" for each of four inequalities viz f(x)<0, f(x) ≤ 0, f(x) > 0 and f(x) ≥ 0.
Solution : Here, given function is a linear function. At y=0,

$$f(x) = 2x + 2 = 0$$

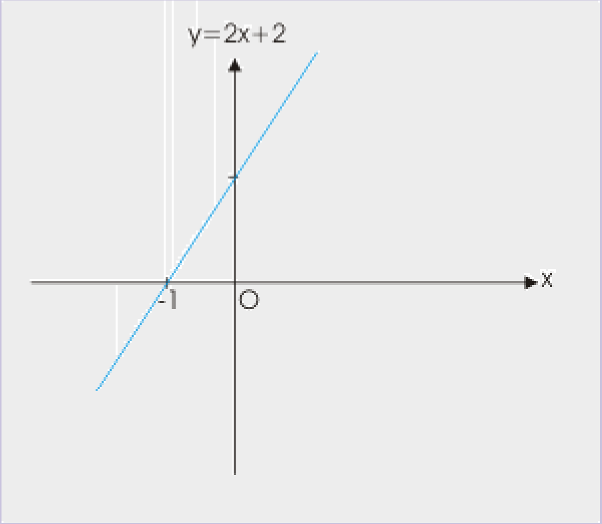
$$\Rightarrow x = -1$$

At x=0,

$$f(x) = 2$$

We draw a line passing through these two points as shown in the figure. From the figure, we conclude that :

Graph of linear function



Graph is continuous for all values of x.

$$\Rightarrow f(x) < 0; \quad x \in (-\infty, -1)$$

$$\Rightarrow f(x) \leq 0; \quad x \in (-\infty, -1]$$

$$\Rightarrow f(x) > 0; \quad x \in (-1, \infty)$$

$$\Rightarrow f(x) \geq 0; \quad x \in [-1, \infty)$$

Polynomial function

A real polynomial, simply referred as polynomial in our study, is an algebraic expression having terms of “x” raised to non-negative numbers, separated by “+” or “-“ sign. A polynomial in one variable is called a univariate polynomial, a polynomial in more than one variable is called a multivariate polynomial. A real polynomial function in one variable is an algebraic expression having terms of real variable “x” raised to non-negative numbers. The general form of representation is :

$$f(x) = a_o + a_1 x + a_2 x^2 + \ldots\ldots + a_n x^n$$

or

$$f(x) = a_o x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots\ldots + a_n$$

Here, a_o ,a_1 ,….,a_n are real numbers. For real function, “x” is real variable and “n” is a non-negative number. An expression like $2x^2 + 2$ is a valid polynomial in “x”. But, $x + 1/x$ is not as $1/x = x^{-1}$ has negative integer power. Also, $3x^{1.2} + 2x$ is not a polynomial as it contains a term with fractional power. Sum and difference of two real polynomials is also a polynomial. Polynomials are continuous function. Its domain is real number set R, whereas its range is either real number set R or its subset. Derivative and anti-derivative (indefinite integral) of a polynomial are also real polynomials.

Degree of polynomial function/ expression

Highest power in the expression is the degree of the polynomial. The degree of the polynomial $x^3 + x^2 + 3$ is 3. The degree “1” corresponds to linear, degree “2” to quadratic, “3” to cubic and “4” to bi-quadratic polynomial. The general form of quadratic equation is :

$$ax^2 + bx + c; \quad a, b, c \in R; \quad a \neq 0$$

Note that “a” can not be zero because degree of function/ expression reduces to 1. Extending this requirement for maintaining order of polynomial, we define polynomial of order “n” as :

$$f(x) = a_o x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots\ldots + a_n; \quad a_0 \neq 0$$

Polynomial equation

The polynomial equation is formed by equating polynomial to zero.

$$f(x) = a_o x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots\ldots + a_n = 0$$

A quadratic equation has the form :

$$f(x) = ax^2 + bx + c = 0$$

The roots of a polynomial equation are the values of “x” for which value of polynomial f(x) becomes zero. If f(a) = 0, then "x=a" is the root of the polynomial. A polynomial equation of degree “n” has at the most “n” roots – real or imaginary. Important point to underline here is that a real polynomial can have imaginary roots.

Solution of polynomial equation is intersection(s) of two equations :

$$y = a_o x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots\ldots + a_n = 0$$

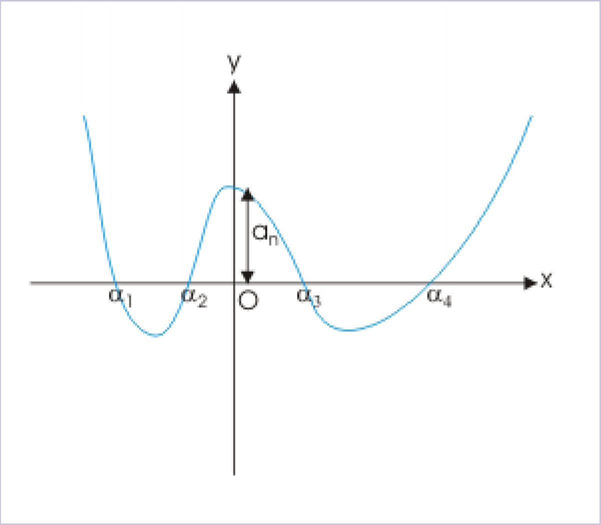
and

$$y = 0 \quad \text{(x-axis)}$$

The solutions of equations (real or complex) are the roots of the polynomial equation. If we plot y=f(x) .vs. y=0 plot, then real roots are x-coordinates (x-intercepts) where plot intersect x-axis. Clearly, graph of polynomial can at most intersect x-axis at “n” points, where “n” is the degree of polynomial. On the other hand, y-intercept of a polynomial is obtained by putting x=0,

$$y = a_0 X0 + a_1 X0 + a_2 X0 + \ldots\ldots + a_n = a_n$$

x and y intercepts of polynomial



Graph of polynomial can at most intersect x-axis at “n” points, where “n” is the degree of polynomial.

Polynomial equation

Some useful deductions about roots of a polynomial equation and their nature are :

1 : A polynomial equation of order n can have n roots – real or imaginary.

2 : Imaginary roots occur in pairs like 1+3i and 1-3i

3 : Roots having square root term occur in pairs 1+√3 and 1-√3.

4 : If a polynomial equation involves only even powers of x and all terms are positive, then all roots of polynomial equation are imaginary (complex). For example, roots of the quadratic equation given here are complex.

$$x^4 + 2x^2 + 4 = 0$$

Descartes rules of signs

Descartes rules are :

(i) Maximum number of positive real roots of a polynomial equation f(x) is equal to number of sign changes in f(x).

(ii) Maximum number of negative real roots of a polynomial equation f(x) is equal to number of sign changes in f(-x).

The signs of the terms of polynomial equation $f(x) = x^3 + 3x^2 - 12x + 3 = 0$ are “+ + - +”. There are two sign changes as we move from left to right. Hence, this cubic polynomial can have at most 2 positive real roots. Further, corresponding $f(-x) = -x^3 + 3x^2 + 12x + 3 = 0$ has signs of term given as “- + + +“. There is one sign change involved here. It means that polynomial equation can have at most one negative root.

Polynomials

Zero polynomial

The function is defined as :

$$y = \mathrm{f}(\mathrm{x}) = 0$$

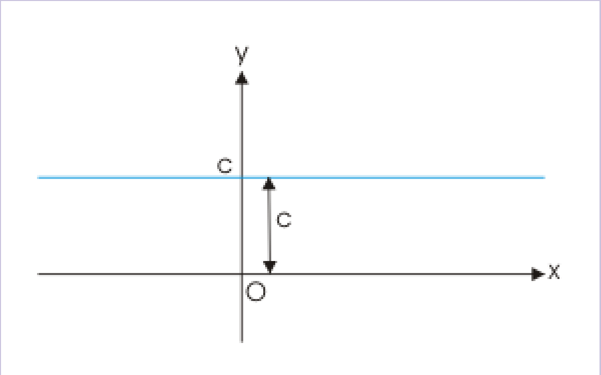
The polynomial “0”, which has no term at all, is called zero polynomial. The graph of zero polynomial is x-axis itself. Clearly, domain is real number set R, whereas range is a singleton set {0}.

Constant function

It is a polynomial of degree 0. The value of constant function is constant irrespective of values of "x". The image of the constant function (y) is constant for all values of pre-images (x).

$$y = \mathrm{f}(\mathrm{x}) = c$$

Constant function



Constant function is a polynomial of degree 0.

The graph of a constant function is a straight line parallel to x-axis. As “y = (f(x) = c” holds for real values of “x”, the domain of constant function is "R". On the other hand, the value of “y” is a single valued constant, hence range of constant function is singleton set {c}.We can treat constant function also as a linear function of the form f(x) = c with m=0. Its graph is a straight line like that of linear function.

There is an interesting aspect about periodicity of constant function. A polynomial function is not periodic in general. A periodic function repeats function values after regular intervals. It is defined as a fuction for which f(x+T) = f(x), where T is the period of the function. In the case of constant function, function value is constant whatever be the value of independent variable. It means that $\mathrm{f}(\mathrm{x} + a_1) = \mathrm{f}(\mathrm{x} + a_2) =$.......... $\mathrm{f}(\mathrm{x}) = c$. Clearly, it meets the requirement with the difference that there is no definite or fixed period like "T". The relation of periodicity, however, holds for any change to x. We, therefore, summarize (it is also the accepted position) that constant function is a periodic function with no period.

Linear function

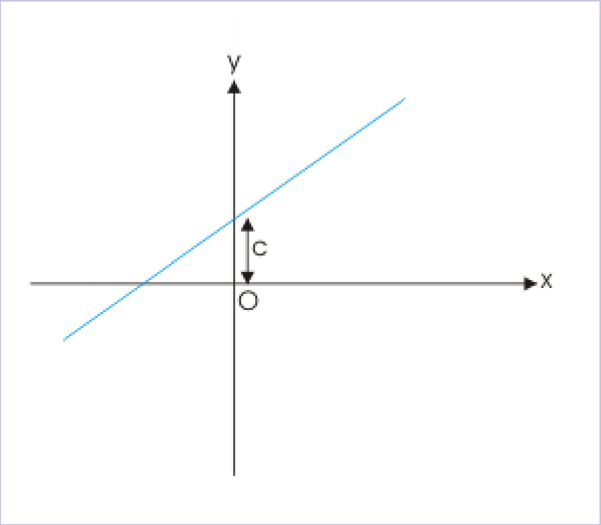
Linear function is a polynomial of order 1.

$$f(x) = a_0x + a_1$$

It is also expressed as :

$$f(x) = mx + c$$

Linear function



Linear function is a polynomial of degree 1.

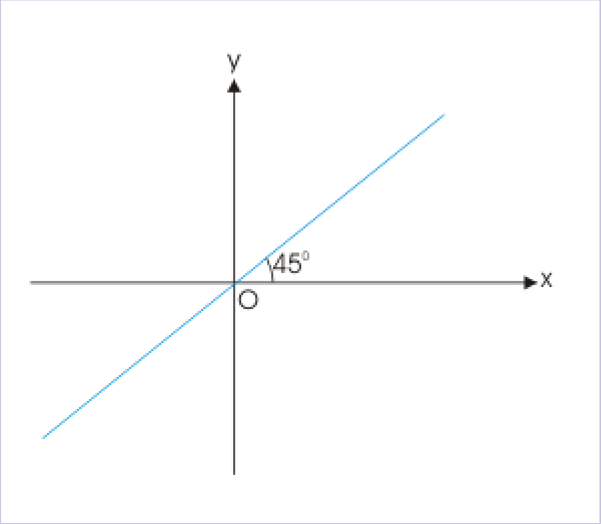
The graph of a linear function is a straight line. The coefficient of “x” i.e. m is slope of the line and c is y-intercept, which is obtained for x = 0 such that f(0) = c. It is clear from the graph that its domain and range both are real number set R.

Identity function

The dependent (y) and independent (x) variables have same value. Identity function is similar in concept to that of identity relation which consists of relation of an element of a set with itself. It is a linear function in which m=1 and c=0. Identity function form is represented as :

$$y = \mathrm{f(x)} = x$$

Identity function



Identity function is a polynomial of degree 1.

The graph of identity function is a straight line bisecting first and third quadrants of coordinate system. Note that slope of straight line is 45°. It is clear from the graph that its domain and range both are real number set R.

Quadratic function

The general form of quadratic function is :

$$\mathrm{f(x)} = ax^2 + bx + c; \quad a, b, c \in R; \quad a \neq 0$$

We shall discuss quadratic function in detail in a separate module and hence discussion of this function is not taken up here.

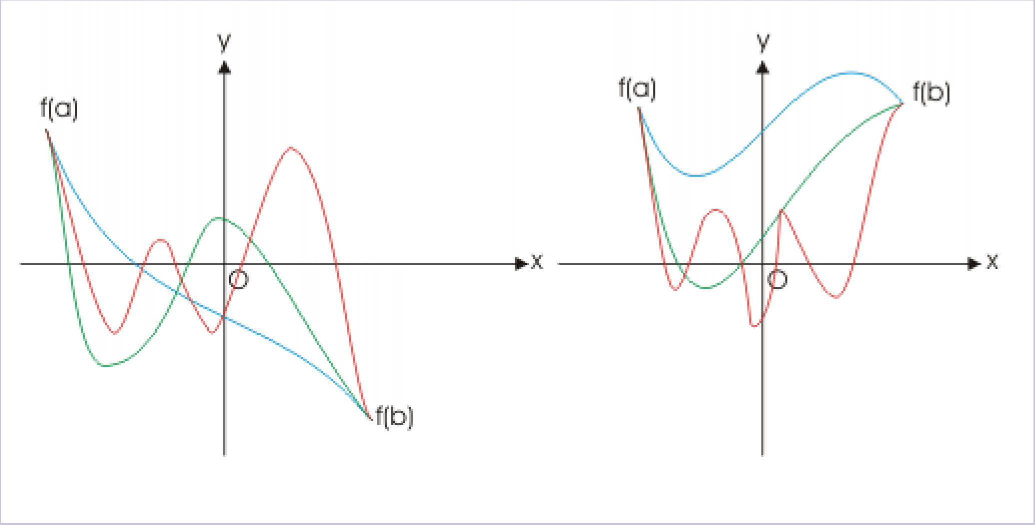
Graph of polynomial function

Graph of polynomial is continuous and non-periodic. If degree is greater than 1, then it is a non-linear graph. Polynomial graphs are analyzed with the help of function properties like intercepts, slopes, concavity, and end behaviors. The may or may not intersect x-axis. This means that it may or may not have real roots. As maximum number of roots of a polynomial is at the most equal to the order of polynomial, we can deduce that graph can at the most intersect x-axis “n” times as maximum numbers of real roots are “n”.

The fact that graph of polynomial is continuous suggests two interesting inferences :

1: If there are two values of polynomial f(a) and f(b) such that f(a)f(b) < 0, then there are at least 1 or an odd numbers of real roots between a and b. The condition f(a)f(b) < 0 means that function values f(a) and f(b) lie on the opposite sides of x-axis. Since graph is continuous, it is bound to cross x-axis at least once or odd times. As such, there are at least 1 or odd numbers of real roots (as shown in the left figure down).

Roots of polynomial function

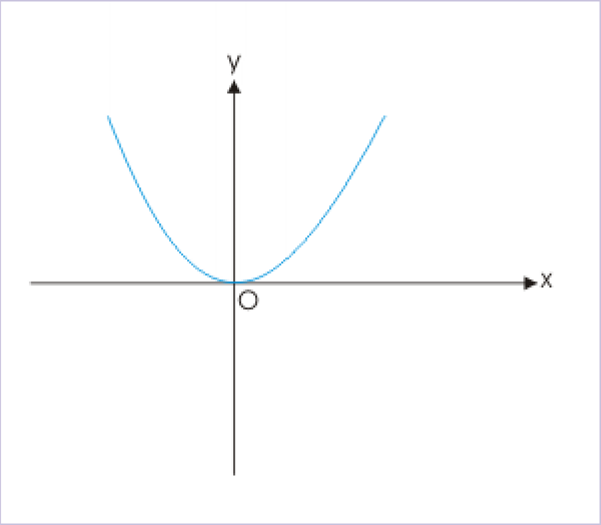


The numbers of x-intercepts depend on nature of product given by f(a)f(b).

2 : If there are two values of polynomial f(a) and f(b) such that f(a)f(b) > 0, then there are either no real roots or there are even numbers of real roots between a and b. The condition f(a)f(b) > 0 means that function values f(a) and f(b) are either both negative or both positive i.e. they lie on the same side of x - axis. Since graph is continuous, it may not cross at all or may cross x-axis even times (as shown in the right figure above). Clearly, there is either no real root or there are even numbers of real roots.

We shall study graphs of quadratic polynomials in a separate module. Further, other graphs will be discussed in appropriate context, while discussing a particular function. Here, we present two monomial quadratic graphs $y = x^2$ and $y = x^3$. These graphs are important from the point of view of generalizing graphs of these particular polynomial structure. The nature of graphs $y = x^n$, where “n” is even integer greater than equal to 2, is similar to the graph of $y = x^2$. We should emphasize that the shape of curve simply generalizes the nature of graph – we need to draw them actually, if we want to draw graph of a particular monomial function. However, we shall find that these generalizations about nature of curve lets us know a great deal about the monomial polynomial. In particular, we can conclude that their domain and range are real number set R.

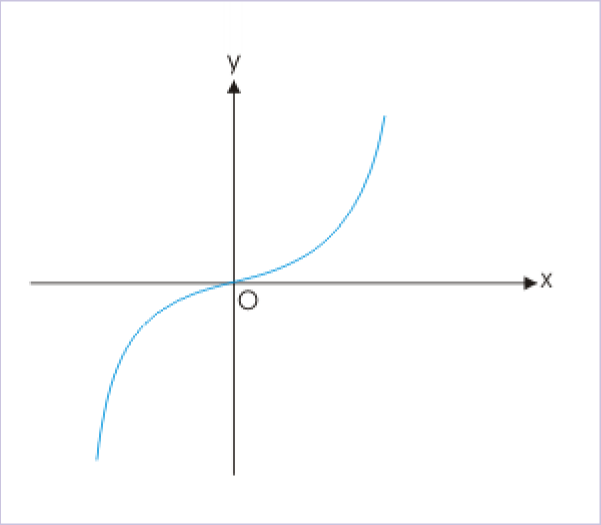
Even degree function



The nature of graphs of degree of positive even integer are similar to the graph shown.

Similarly, the nature of graphs $y = x^n$, where “n” is odd number integer greater than 2, is similar to the graph of $y = x^3$.

Odd degree function



The nature of graphs of degree of positive odd integer are similar to the graph of shown.

Quadratic polynomial function

We are already acquainted with quadratic equation and its roots. In this module, we shall study quadratic expression from the point of view of a function. It is a polynomial function of degree 2. The general form of quadratic expression/ function is :

$$f(x) = ax^2 + bx + c; \quad a, b, c \in R, \quad a > 0$$

Elements of quadratic equation

Quadratic equation

Quadratic equation is obtained by equating quadratic function to zero. General form of quadratic equation corresponding to quadratic function is :

$$ax^2 + bx + c = 0; \quad a, b, c \in R, \quad a > 0$$

Discriminant of quadratic equation

Nature of a given quadratic function is best understood in terms of discriminant, D, of corresponding quadratic equation. This is given as :

$$D = b^2 - 4ac$$

Roots of quadratic equation

Quadratic equation is obtained by equating quadratic function to zero. Quadratic equation has at most two roots. The roots are given by :

$$\alpha = \frac{-b - \sqrt{D}}{2a} = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

$$\beta = \frac{-b + \sqrt{D}}{2a} = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

Properties of roots of quadratic equation

1 : If D>0, then roots are real and distinct.

2 : If D=0, then roots are real and equal.

3 : If D<0, then roots are complex conjugates with non-zero imaginary part.

4 : If D>0; a,b,c∈T (rational numbers) and D is a perfect square, then roots are rational.

5 : If D>0; a,b,c∈T (rational numbers) and D is not a perfect square, then roots are radical conjugates.

6 : If D>0; a=1;b,c∈Z (integer numbers) and roots are rational, then roots are integers.

7 : If a quadratic equation has more than two roots, then the function is an identity in x and a=b=c=0.

8 : If a quadratic equation has one real root and a,b,c∈R, then other root is also real.

Elements of quadratic function

Zeroes of quadratic function

The real roots of the quadratic equation are zeroes of quadratic function. The zeroes of quadratic function are real values of x for which value of quadratic function becomes zero. On graph, zeros are the points at which graph intersects y=0 i.e. x-axis.

Graph of quadratic function

Graph reveals important characteristics of quadratic function. The graph of quadratic function is a parabola. Working with the quadratic function, we have :

$$y = ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right)$$

In order to complete square, we add and subtract $b^2/4a^2$ as :

$$\Rightarrow y = a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} + \frac{c}{a} - \frac{b^2}{4a^2}\right)$$

$$\Rightarrow y = a\left\{\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}\right\}$$

$$\Rightarrow y + \frac{b^2 - 4ac}{4a} = a\left(x + \frac{b}{2a}\right)^2$$

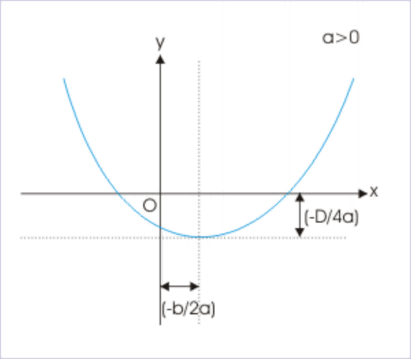
$$\Rightarrow y + \frac{D}{4a} = a\left(x + \frac{b}{2a}\right)^2$$

$$\Rightarrow Y = aX^2$$

Where,

$$X = x + \frac{b}{2a} \quad \text{and} \quad Y = y + \frac{D}{4a}$$

Graph of quadratic function



The graph is parabola.

Clearly, $Y = aX^2$ is an equation of parabola having its vertex given by (-b/2a, -D/4a). When a>0, parabola opens up and when a<0, parabola opens down. Further, parabola is symmetric about x=-b/2a.

Maximum and minimum values of quadratic function

The graph of quadratic function extends on either sides of x-axis. Its domain, therefore, is R. On the other hand, value of function extends from vertex to either positive or negative infinity, depending on whether "a" is positive or negative.

When a > 0, the graph of quadratic function is parabola opening up. The minimum and maximum values of the function are given by :

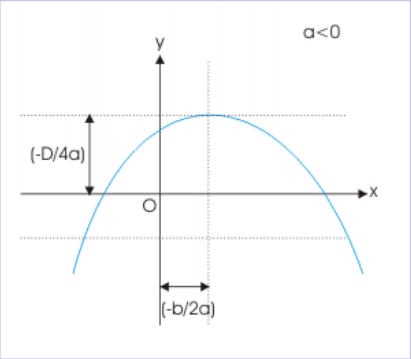
$$y_{\text{min}} = -\frac{D}{4a} \quad \text{at} \quad x = -\frac{b}{2a}$$

$$y_{\text{max}} \Rightarrow \infty$$

Clearly, range of the function is [-D/4a, ∞).

When a < 0, the graph of quadratic function is parabola opening down. The maximum and minimum values of the function are given by :

Graph of quadratic function



The graph is parabola, which opens down.

$$y_{\text{max}} = -\frac{D}{4a} \quad \text{at} \quad x = -\frac{b}{2a}$$

$$y_{\text{min}} \Rightarrow -\infty$$

Clearly, range of the function is (-∞, -D/4a].

Example:

Problem : Determine range of $f(x) = -3x^2 + 2x - 4$

Solution : The determinant of corresponding quadratic equation is :

$$D = b^2 - 4ac = 4 - 4X(-3)X(-4) = 4 - 48 = -44 \quad \Rightarrow D < 0$$

$$a = -3 \quad \Rightarrow a < 0$$

The graph of function is parabola opening down. Its vertex represents the maximum function value. The maximum and minimum values of function are given by :

$$\Rightarrow y_{\max} = -\frac{D}{4a} = -\frac{-44}{4X-3} = -\frac{44}{12} = -\frac{11}{3}$$

$$y_{\min} \Rightarrow -\infty$$

Range = (-∞, -11/3)

Nature of quadratic function

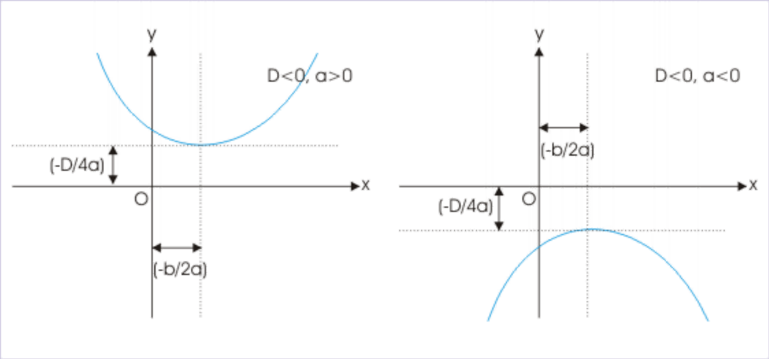
The discriminant of corresponding quadratic equation and coefficient of term “x^2 ” of quadratic function together determine nature of quadratic function and hence its graph. Graphs of quadratic function is intuitive and helpful to remember results. As a matter of fact, we can interpret all properties of quadratic function, if we can draw its graph.

Case 1 : D<0

If D<0, then roots are complex conjugates. It means graph of function does not intersect x-axis. If a > 0, then parabola opens up. The value of quadratic function is positive for all values of x i.e.

$$D < 0, a > 0 \quad \Rightarrow f(x) > 0 \quad \text{for} \quad x \in R$$

Graph of quadratic function



The discriminant is negative.

If a < 0, then parabola opens down. The value of quadratic function is negative for all values of x i.e.

$$D < 0, a < 0 \quad \Rightarrow f(x) < 0 \quad \text{for} \quad x \in R$$

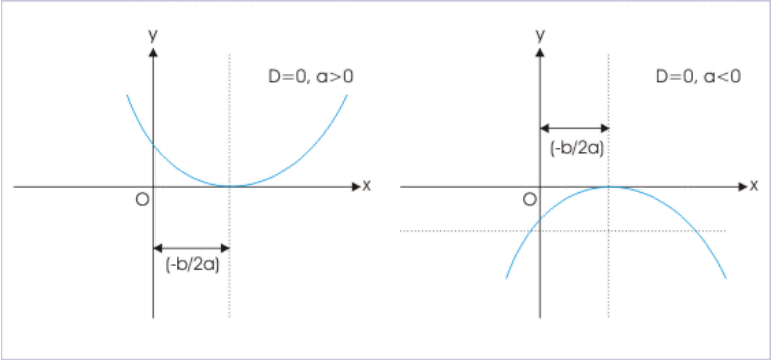
Sign rule : If D<0, then sign of function is same as that of “a” for all values of x in R.

Case 2 : D=0

If D=0, then roots are equal and is given by –b/2a. It means graph of function just touches x-axis. If a > 0, then parabola opens up. The value of quadratic function is non-negative for all values of x i.e.

$$D = 0, a > 0 \quad \Rightarrow f(x) \geq 0 \quad \text{for} \quad x \in R$$

Graph of quadratic function



The discriminant is zero.

If a < 0, then parabola opens down. The value of quadratic function is non-positive for all values of x i.e.

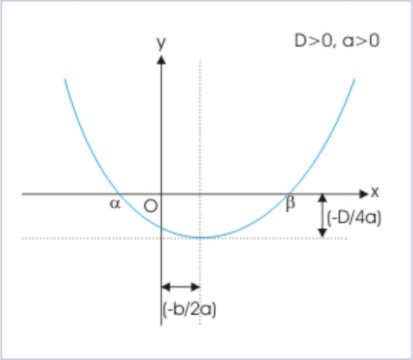
$$D = 0, a < 0 \quad \Rightarrow f(x) \leq 0 \quad \text{for} \quad x \in R$$

Sign rule : If D=0, then sign of function is same as that of "a" for all values of x in R except at x=-b/2a, at which f(x)=0. We do not associate sign with zero.

Case 3 : D>0

If D>0, then roots are unequal and are given by (–b±D)/2a. It means graph of function intersects x-axis at α and β (β>α). If a > 0, then parabola opens up. The value of quadratic function is positive for all values of x in the interval (-∞,α) U (β,∞).The values of quadratic function are zero for values of x ∈{α,β}. The value of quadratic function is negative for all values of x in the interval (α,β).

Graph of quadratic function



The discriminant is positive and a is positive.

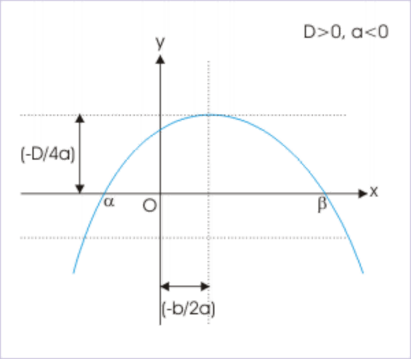
$$D > 0,a > 0 \quad \Rightarrow f(x) > 0 \quad \text{for} \quad x \in (-\infty, \alpha) \cup (\beta, \infty) \quad \Rightarrow \text{Sign of function same as that of “a”}$$

$$D > 0,a > 0 \quad \Rightarrow f(x) = 0 \quad \text{for} \quad x \in \{\alpha, \beta\}$$

$$D > 0,a > 0 \quad \Rightarrow f(x) < 0 \quad \text{for} \quad x \in (\alpha, \beta) \quad \Rightarrow \text{Sign of function opposite to that of “a”}$$

If a < 0, then parabola opens down. The value of quadratic function is positive for all values of x in the interval (α,β).The values of quadratic function are zero for values of x ∈{α,β}. The value of quadratic function is negative for all values of x in the interval (-∞,α) U (β,∞).

Graph of quadratic function



The discriminant is positive and a is negative.

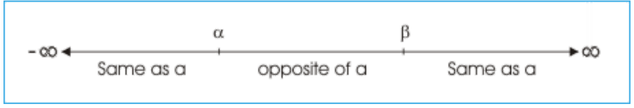
$$D > 0,a < 0 \quad \Rightarrow f(x) < 0 \quad \text{for} \quad x \in (\alpha, \beta) \quad \Rightarrow \text{Sign of function same as that of “a”}$$

$$D > 0,a < 0 \quad \Rightarrow f(x) = 0 \quad \text{for} \quad x \in \{\alpha, \beta\}$$

$$D > 0,a < 0 \quad \Rightarrow f(x) > 0 \quad \text{for} \quad x \in (-\infty, \alpha) \cup (\beta, \infty) \quad \Rightarrow \text{Sign of function opposite to that of “a”}$$

Sign rule : If D>0, then domain of function, which is R, is divided at root points in three intervals. The signs of function in side intervals are same as that of “a”, whereas sign of function in the middle interval is opposite to that of “a”.

Sign rule of quadratic function



Signs of function.

Examples

Example:
Problem : Determine interval of “a” for which graph of $x^2 + (a-1)x + 16$ lie above x-axis.

Solution : Here coefficient of "x^2 " is positive. Now, the graph of quadratic function lie above x-axis when D>0 and a>0.

$$D = (a-1)^2 - 4X1X16 < 0 \quad \Rightarrow (a-1)^2 - 64 < 0$$

$$\Rightarrow (a-1)^2 - 8^2 < 0 \quad \Rightarrow (a-1+8)(a-1-8) < 0$$

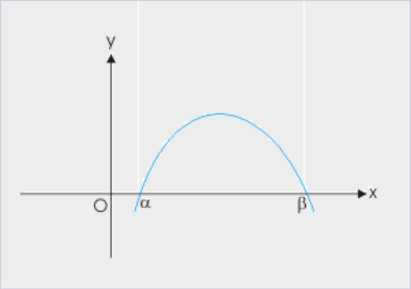
$$\Rightarrow (a+7)(a-9) < 0$$

$$\Rightarrow -7 < a < 9$$

Example:

Problem : The graph of a quadratic expression $f(x) = ax^2 + bx + c$ is shown in the figure. Determine signs of a,b,c.

Graph of quadratic function



The graph is a parabola.

Solution : The parabola opens downward. It means a<0. Since both roots are positive, their sum is also positive.

$$\alpha + \beta = -\frac{b}{a} > 0 \quad \Rightarrow \text{Signs of a and b are opposite.}$$

It means b>0. Now, putting x=0 in the function, we have value of function as :

$$f(0) = aX0^2 + bX0 + c = c$$

From figure, graph can intersects y-axis only at negative y-value. Hence, "c" is negative.

Quadratic inequality

Value of quadratic function is positive, zero or negative depending on the nature of coefficient of "x^2 " and discriminant, D. Accordingly, a quadratic function will hold true for particular interval(s) depending upon these parameters. Consider an example :

$$\Rightarrow 2x^2 - 5x - 3 > 0$$

$$\Rightarrow 2x^2 - 6x + x - 3 > 0$$

$$\Rightarrow 2x(x-3) + (x-3) > 0$$

$$\Rightarrow (2x+1)(x-3) > 0$$

$$\Rightarrow x = -\frac{1}{2}, 3$$

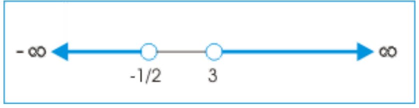
Here,

$$a = 2 \quad \Rightarrow a > 0$$

$$D = 25 - 4X2X3 = 1 \quad \Rightarrow D > 0$$

Under these conditions, given quadratic function is positive for $(-\infty, -1/2) \cup (3{,}\infty)$. On the real number line, the intervals are shown as :

Quadratic inequality



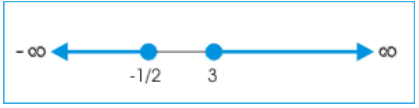
Interval in which f(x) is positive.

Let us consider slightly changed inequality involving less than equal sign,

$$2x^2 - 5x - 3 \leq 0$$

Again quadratic function is positive for $(-\infty, -1/2) \cup (3{,}\infty)$. However, equality is also allowed. It means valid intervals should also include root points. The modified valid interval corresponding to "less than equal to inequality" is : $(-\infty, -1/2] \cup [3{,}\infty)$.

Quadratic inequality



Interval in which f(x) is non-negative.

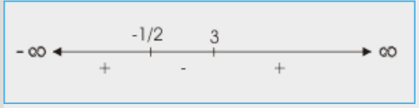
From this illustration, it is clear that we can determine valid interval(s) of x, provided we know the signs of quadratic function in different intervals. If equality is also allowed as in the case of “less than equal to” or “greater then equal to”, then we need to include root points also.

Example:

Problem : Determine the interval of x for which $f(x) = 2x^2 - 5x - 3$ is non-positive and negative.

Solution : As already determined earlier, roots of function are -1/2, 3. Also a> 0 and D > 0. Sign rule for the function is shown here :

Sign rule of quadratic inequality



Signs of function.

From the figure, it is clear that function is non-positive i.e. f(x) ≤0 in the interval [-1/2, 3]. Also, function is negative i.e. f(x)<0 in the interval (-1/2, 3).

Example:

Problem : A quadratic function is given by $f(x) = x^2 - 4x + 4$. Find solution for each of four inequalities viz f(x)<0, f(x) ≤ 0, f(x) > 0 and f(x) ≥ 0.

Solution : Here, coefficient of "x^2 " is 1. Thus,

$$\Rightarrow a > 0$$

Determinant of corresponding quadratic equation is :

$$D = (-4)^2 - 4X1X4 = 0$$

For D=0 and a>0, f(x) ≥0. Further, since D=0, it means that corresponding quadratic equation has one real root. Now, the root is :

$$\alpha = -\frac{b}{2a} = -\frac{-4}{2X1} = 2$$

This means that f(x) is positive for all values of x except for x=2. At x=2, f(x)=0. It follows then that :

$$f(x) \geq 0; \quad x \in R$$

$$f(x) > 0; \quad x \in R - \{2\}$$

$$f(x) \leq 0; \quad x \in \{2\}$$

$$f(x) < 0; \quad \text{No solution}$$

Exercise:

Problem:

A quadratic function is given by $f(x) = -x^2 + 2x - 4$. Find solution for each of four inequalities viz f(x)<0, f(x) ≤ 0, f(x) > 0 and f(x) ≥ 0.

Solution:

Here, coefficient of "x^2 " is -1. Thus,

$$\Rightarrow a < 0$$

Determinant of corresponding quadratic equation is :

$$D = (2)^2 - 4X - 1X - 4 = -8 \quad \Rightarrow D < 0$$

For D<0 and a<0, f(x) < 0. Further, since D<0, it means that corresponding quadratic equation has no real root. This means that f(x) is negative for all values of x. It follows then that

$$\begin{aligned} f(x) < 0; &\quad x \in R \\ f(x) \leq 0; &\quad x \in R \\ f(x) > 0; &\quad \text{No solution} \\ f(x) \geq 0; &\quad \text{No solution} \end{aligned}$$

Acknowledgment

Author wishes to thank Mr. Ralph David for helping to correct a mistake in the module.

Rational function

Rational function is defined in similar fashion as rational number is defined in terms of numerator and denominator. Implicitly, we refer “real” rational function here. It is defined as the ratio of two real polynomials with the condition that polynomial in the denominator is not a zero polynomial.

$$f(x) = \frac{p(x)}{q(x)}; \quad q(x) \neq 0$$

Rational function is not defined for values of x for which denominator polynomial evaluates to zero as ratio “p(x)/0” is not defined. Some examples of rational function are :

$$f(x) = \frac{2x^2 - x + 1}{2x^2 - 5x - 3}; \quad x \neq -\frac{1}{2}, x \neq 3$$

$$g(x) = \frac{x + 1}{2x^2 - x + 1}$$

$$h(x) = \frac{2x^4 - x^2 + 1}{x + 1}; \quad x \neq -1$$

Note second example function, g(x) above. There is no exclusion point for this rational polynomial. The denominator polynomial is $2x^2 - x + 1$, whose determinant is negative and coefficient of x^2 term is positive. It means denominator of g(x) is positive for all values of x. We should also note that values of x being excluded are points - not a continuous interval. Further, the notation to denote exclusion is an “inequation” – not “inequality” – because notation $x \neq -1$ negates corresponding equation x = -1. Recall that inequality, on the other hand, compares relative values.

Domain of rational function

Domain of rational function is domain of numerator polynomial minus exclusion points as determined by zeroes of denominator polynomial. Since domain of polynomial is R, domain of rational polynomial is R minus exclusion points determined by denominator. The domains for three rational functions given above are :

$$\text{Domain of f(x)} = R - \left\{-\frac{1}{2}, 3\right\}$$

$$\text{Domain of g(x)} = R$$

$$\text{Domain of h(x)} = R - \{-1\}$$

Important properties of rational function

Important properties are :

- Singularity or exception point
- Holes
- Asymptotes – vertical, horizontal and slant
- x and y intercepts

Singularities

Singularities are x-values for which denominator of rational function is zero. The function is not defined for such x-values and as such these values are excluded from the domain set of the function. These points are also called exception points. Function is not defined at these points.

Factorizing numerator and denominator of rational function helps to identify singularities of algebraic rational function. Singularities correspond to x values resulting from equating linear factors in denominator to zero. The important thing to note here is that singularity or exception occurs when denominator of rational function turns zero – no matter whether linear factor in the denominator cancels out with the linear factor in numerator or not. To understand this point, let us consider few rational functions given below :

$$f(x) = \frac{(x-1)(x+2)}{(x-1)(x+1)}$$

$$g(x) = \frac{(x-1)^2(x+2)}{(x-1)(x+1)}$$

$$h(x) = \frac{(x-1)(x+2)}{(x-1)^2(x+1)}$$

We can see that h(x) contains a linear factor (x-1) in the denominator after cancellation of like linear factors. On the other hand, functions f(x) and g(x) do not contain (x-1) in the denominator after cancellation of like linear factors. The function g(x), however, contains (x-1) in the numerator after cancellation. Notwithstanding these possibilities, denominator of the rational function turns zero at x=1. As such, the point specified by x=1 is singularity for all three function forms shown above.

We shall see that either a hole or vertical asymptote occurs at the point of exception i.e. singularity. It depends on how do the linear factors in the denominator relate to linear factors in the numerator.

Exercise:

Problem: Find singularities of function given by :

$$\frac{x^2-3x+2}{x^2-2x-3}$$

Solution:

Factorizing into linear factors, we have :

$$\frac{x^2-3x+2}{x^2-2x-3} = \frac{(x-1)(x-2)}{(x+1)(x-3)}$$

Equating denominator to zero, we have :

$$\Rightarrow (x+1)(x-3)=0$$

$$\Rightarrow x=-1 \quad \text{or} \quad 3$$

Thus, singularities are -1 and 3.

Holes

Hole exists at a singularity when corresponding linear factor of the denominator cancel out completely or when linear factor remains in the numerator after cancellation. Hole is a point on the graph where function value is not defined. It is a point having x and y coordinates. We determine x-coordinate by equating linear factor in denominator to zero. Its y-coordinate is obtained by plugging x-value in the reduced (after cancellation of common factors in numerator and denominator) function form. As pointed out, there are two situations with respect to position of a hole :

1: If linear factor cancels completely, then hole lies any where but not on x-axis.

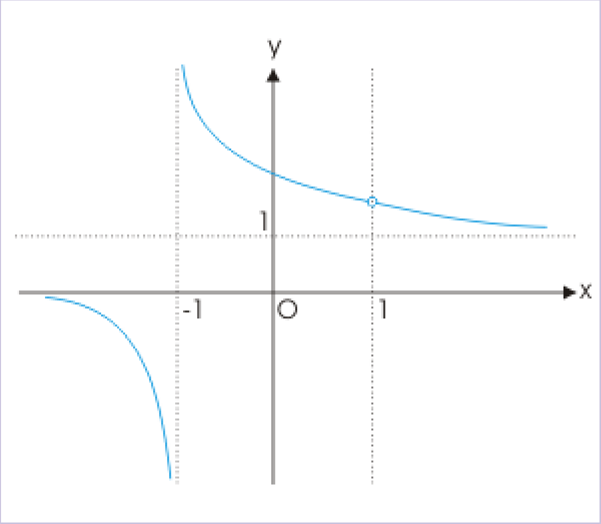
$$f(x)=\frac{(x-1)(x+2)}{(x-1)(x+1)}$$

Here, singularities occur at x= -1 and 1. The linear factor (x-1) is present in both numerator and denominator and as such cancels out completely. Therefore, there is a hole at x=1. On the other hand, no cancellation is involved at x=-1. There exists a vertical asymptote at x=-1. Important point to realize here is that if linear factor is present in denominator only and there is no cancellation involved, then x-value corresponding to linear factor is not the x-coordinate of hole. We shall learn about vertical asymptote subsequently. Now, the y-coordinate of hole is :

$$\Rightarrow f(x)=\frac{(x+2)}{(x+1)}=\frac{3}{2}=1.5$$

The graph of function, f(x), is shown in the figure below :

Rational function



Singularities, holes and vertical asymptote

2: If linear factor remains in the numerator after cancellation, then hole lies on x-axis. The graph tends to intercept x-axis. As such, hole exists at x-axis.

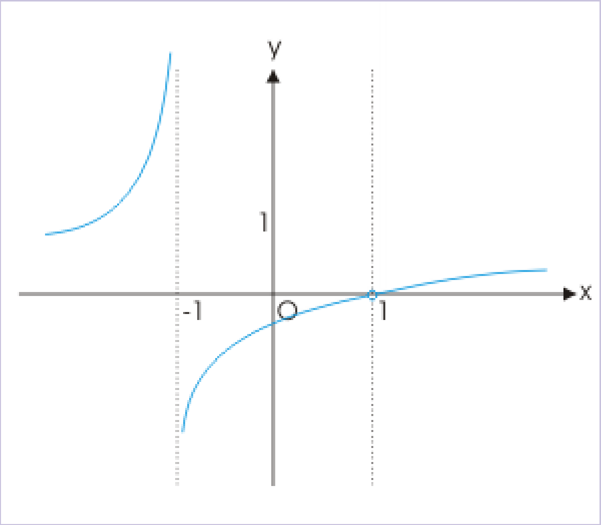
$$g(x) = \frac{(x-1)^2(x+2)}{(x-1)(x+1)}$$

Here, singularities occur at x= -1 and 1. There is a vertical asymptote at x=-1, but a hole at x=1. The y-coordinate of hole is :

$$\Rightarrow g(x) = \frac{(x-1)(x+2)}{(x+1)} = \frac{0X3}{2} = 0$$

Thus, hole lies on x-axis.

Rational function



Singularities, holes and vertical asymptote

Asymptotes

An asymptote is a straight line at singularity which graph of function tends to approach but never touches. The difference between graph and asymptotes is infinitesimally small as the graph is extended away from x-axis. Important to note is that graph neither touches or crosses the asymptote. In the case of vertical asymptote, the function values tend to be either positive or negative large number or a combination of two on either side of the vertical asymptote. In the case of horizontal asymptote, the function values tend to be a finite value.

Vertical asymptote

Vertical asymptote is a vertical line including y-axis to which graph of function comes closer and closer but never touches. Vertical asymptotes correspond to very large y-values, where difference between x-value and asymptote is infinitesimally small. An equation of vertical asymptote has the form,

$$x = c$$

Vertical asymptotes occur at singularity when linear factor in the denominator remains after cancellation or otherwise. Let us investigate three functions given earlier for existence of vertical asymptote.

$$f(x) = \frac{(x-1)(x+2)}{(x-1)(x+1)}$$

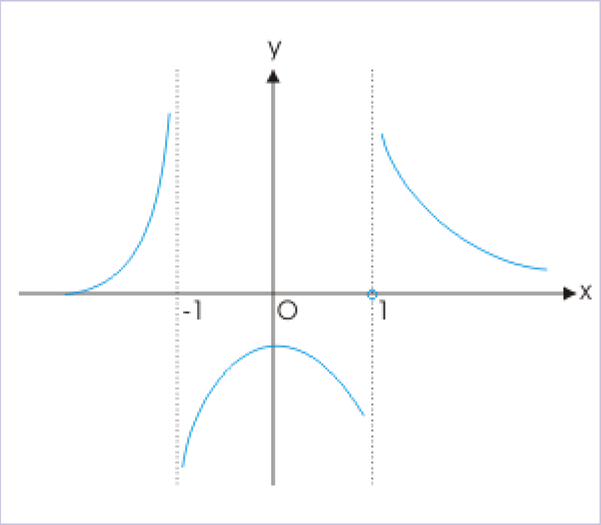
$$g(x) = \frac{(x-1)^2(x+2)}{(x-1)(x+1)}$$

$$h(x) = \frac{(x-1)(x+2)}{(x-1)^2(x+1)}$$

For function, f(x), singularities exists at x=1 and -1. Here x=1 is a hole as linear factor (x-1) cancels out completely. The linear factor (x+1), however, does not cancel out. Thus, a vertical asymptote exists at x=-1. See graph shown earlier drawn for f(x).

For function, g(x), also singularities exists at x=1 and -1. Here x=1 is a hole on x-axis as linear factor (x-1) remains in the numerator after cancellation. The linear factor (x+1), however, does not cancel out. Thus, a vertical asymptote exists at x=-1. See graph shown earlier drawn for g(x).

For function, h(x), also singularities exists at x=1 and -1. Here, a vertical asymptote exists at x=1 as linear factor (x-1) remains in the denominator. The linear factor (x+1), however, does not cancel out. Thus, a vertical asymptote also exists at x=-1. See graph shown here for h(x).
Rational function



Singularities, holes and vertical asymptote

Nature of vertical asymptote

The function value assumes large values close to singularity where asymptote exists. The values are directed either in the same of opposite directions. It depends on the polarity of reduced function. If the reduced function has linear factor raised to even power, then values asymptotes in the same direction. On the other hand, if the reduced function has linear factor raised to odd power, then values asymptotes in opposite directions. Let us consider function as defined here,

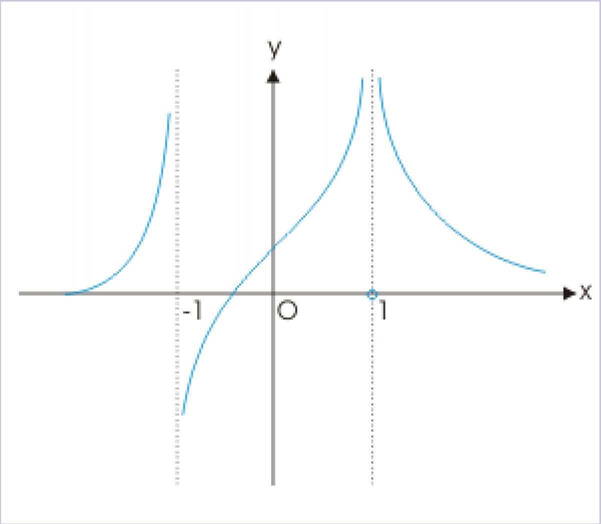
$$f(x) = \frac{(x-1)(x+2)}{(x-1)^3(x+1)}$$

After simplification, the function reduces to :

$$\Rightarrow f(x) = \frac{(x+2)}{(x-1)^2(x+1)}$$

Clearly, (x-1) is raised to even power 2. The graph asymptotes towards large positive values i.e. in the same direction from either side of the asymptote. On the other hand, the linear factor (x+1) is raised to 1 i.e. odd power. Hence function value asymptotes in opposite directions.

Rational function



Asymptotes

Horizontal asymptote

Horizontal asymptote is a horizontal line including x-axis to which graph of function comes closer and closer but never touches. The difference between y-value and asymptote is infinitesimally small for large values of x. An equation of horizontal asymptote has the form,

$$y = c$$

Existence of horizontal asymptote depends on the degree of polynomial in the numerator (n) and degree of polynomial in the denominator (m). There are following three cases :

1: If n>m, then there is no horizontal asymptote. However, if n=m+1, then there exists slant asymptote.

2: If n<m, then x-axis is horizontal asymptote.

3: If n=m, then there is horizontal asymptote exists. In this case, the equation of horizontal asymptote is :

$$y = \frac{\text{Coefficient of highest power term in numeratory}}{\text{Coefficient of highest power term in denominator}}$$

Exercise:

Problem: 1. Find horizontal asymptote :

$$f(x) = \frac{2x^2 + x + 1}{x^2 + 16}$$

Solution:

Here, coefficient of highest power term is 2 in numerator and 1 in denominator. Hence, horizontal asymptote is given by :

$$y = \frac{2}{1} = 2$$

Exercise:

Problem: 2. Find horizontal asymptote :

$$f(x) = \frac{-x(x+1)}{x^4 + 16}$$

Solution:

The order of highest power term is 2 in numerator and 4 in denominator. Thus, n<m. Hence, x-axis is horizontal asymptote.

$$y = 0$$

Exercise:

Problem: 3. Find horizontal asymptote :

$$f(x) = \frac{x\left(x^2 - 1\right)}{x + 2}$$

Solution:

The order of highest power term is 3 in numerator and 1 in denominator. Here, n>m. Hence, there is no horizontal asymptote.

Slant asymptotes

Slant asymptote is a line that the graph approaches. This line is neither vertical nor horizontal. A rational function has a slant or oblique asymptote when order of numerator (n) is greater than order of denominator (m).

The equation of slant asymptote is obtained by dividing numerator polynomial by denominator polynomial. The quotient of division is equation of asymptote. Clearly, asymptote is a straight line. As such, quotient should be a linear expression. The requirement that asymptote is a straight line implies that the order of numerator polynomial is higher than order of denominator polynomial by 1 i.e. n=m+1.

In the nutshell, slant asymptote exists when n=m+1. The slant asymptote is obtained by dividing numerator and denominator. We neglect remainder. The equation of the slant asymptote is given by quotient equated to “y”.

Exercise:

Problem: Find slant asymptote :

$$f(x) = \frac{x^2}{x+1}$$

Solution:

Division here yields quotient as “x-1”. Hence, equation of slant asymptote is :

$$y = x - 1$$

x-intercepts

The x-intercepts are also known as zeroes of function or real roots of corresponding equation when function is equated to zero. Since function is many-one, there can be more than one x-intercept. On graphs, x-intercepts are points on x-axis, where graph intersects it. Thus, x-intercepts are x-values where function value becomes zero.

$$f(x) = 0$$

In the case of rational function, x-intercepts exist when numerator turns zero. In other words, x-intercepts are x-values for which numerator of the function turns zero.

$$g(x) = \frac{f(x)}{h(x)} = 0$$

$$g(x) = 0 \quad \Rightarrow f(x) = 0$$

We determine x-intercepts by solving equation formed by equating function to zero. It is hepful to know that real polynomial of odd degree has a real root and, therefore, at least one x-intercept. In the case of rational function, the function is not defined for values of x when denominator turns zero. It means that there will be no x-intercept corresponding to linear factor which is common to denominator. Consider the function given here :

$$h(x) = \frac{(x-1)(x+2)}{(x-1)^2(x+1)}$$

Equating numerator to zero, we have :

$$\Rightarrow (x-1)(x+2) = 0$$

$$\Rightarrow x = 1, -2$$

But (x-1) is also linear factor in denominator. It means that point x=1 is a singularity. Hence, x-intercept is only x=-2 as function is not defined at x=1.

Exercise:

Problem: Find x-intercepts of reciprocal function :

$$f(x) = \frac{1}{x}$$

Solution:

Here numerator is 1 and can not be zero. Thus, reciprocal function does not have x-intercepts.

Exercise:

Problem: Find x-intercepts of function given by :

$$\frac{x^2 - 3x + 2}{x^2 - 2x - 3}$$

Solution:

Factorizing, we have :

$$\frac{x^2-3x+2}{x^2-2x-3}=\frac{(x-1)(x-2)}{(x+1)(x-3)}$$

Equating numerator to zero, we have :

$$\Rightarrow (x-1)(x-2)=0$$

$$\Rightarrow x=1 \quad \text{or} \quad 2$$

Thus, x-intercepts are 1 and 2.

y-intercepts

This is function or y value when x is zero. Functions are many-one relation. Thus, there can be only one y-intercept. The y-intercept is calculated as :

$$y=f(0)$$

In the case of rational function, we can not determine y-intercept if the function is not defined at x=0.

Exercise:

Problem: Find y-intercepts of reciprocal function :

$$y=\frac{1}{x}$$

Solution:

The equation can not be solved for x=0. Thus, reciprocal function does not have y-intercepts.

Exercise:

Problem: Find y-intercepts :

$$f(x) = 2x - 3$$

Solution:

Putting x=0,

$$f(0) = -3$$

Real values of rational function

In order to find real values of rational function for real x, we rearrange the given rational function to form a quadratic equation in x. Let us consider a rational polynomial of degree 2 given by (we mean that the highest degree of polynomials involved in the ratio is 2) :

$$y = f(x) = \frac{a_1x^2 + b_1x + c_1}{a_2x^2 + b_2x + c_2}$$

Rearranging to form a quadratic equation in x, we have :

$$\Rightarrow y\left(a_2x^2 + b_2x + c_2\right) = a_1x^2 + b_1x + c_1$$

$$\Rightarrow (ya_2 - a_1)x^2 + (yb_2 - b_1)x + c_2y - c_1 = 0$$

For x to be real, D≥0. Hence,

$$\Rightarrow D = (yb_2 - b_1)^2 - 4x(ya_2 - a_1)(c_2y - c_1) \geq 0$$

We see that discriminant itself is a quadratic inequality. Depending on the nature of coefficient of “y^2 ” in the quadratic equation and determinant of the corresponding quadratic equation, the inequality is solved for “y”. This, in turn, allows us to determine the real interval(s) of y corresponding to real x.

We should clearly understand that these are real values of y corresponding to real x. This interval need not be the range of the function. Recall that range of a function contains values of y for values of x in the domain of function – not all real values of x. Now function is not defined for certain values of x, which are zeroes of denominator of the rational function. Hence domain is not the real number set R. This distinction should always be kept in mind.

Example:

Problem : Find value of x for which given function has least value.

$$y = f(x) = \frac{x^2 - 6x + 5}{x^2 + 2x + 1}$$

Solution : Rearranging to form a quadratic equation in x, we have :

$$\Rightarrow yx^2 + 2yx + y = x^2 - 6x + 5$$

$$\Rightarrow (y-1)x^2 + 2(y+3)x + y - 5 = 0$$

For x real, D≥0.

$$\Rightarrow 4(y+3)^2 - 4(y-1)X(y-5) \geq 0$$

$$\Rightarrow \left(y^2 + 6y + 9\right) - \left(y^2 - 6y + 5\right) \geq 0$$

$$\Rightarrow 12y + 4 \geq 0$$

$$\Rightarrow y \geq -\frac{1}{3}$$

The real values of y, therefore, lies in the interval [-1/3, ∞). The least value of y in the interval of real values is -1/3. We should, however, check that value of y=-1/3 does not correspond to value of x which is not permitted. Here, the denominator polynomial is $x^2+2x+1=(x+1)^2$. Thus, $x \ne -1$. The domain of the function is R-{-1}.

Now, we calculate value of x corresponding to y as :

$$\Rightarrow -\frac{1}{3}=\frac{x^2-6x+5}{x^2+2x+1}$$

$$\Rightarrow x^2-4x+4=0$$

$$\Rightarrow (x-2)^2=0$$

$$\Rightarrow x=2$$

This point belongs to the domain of the function as it is different to excluded value. Hence, least value of y is -1/3.

Example:

Problem : For what values of “a”, the function given here assumes all real values for real x.

$$y=f(x)=\frac{ax^2+3x-4}{3x-4x^2+a}$$

Solution : Rearranging to form a quadratic equation in x, we have :

$$\Rightarrow 3yx-4yx^2+ay=ax^2+3x-4$$

$$\Rightarrow (a+4y)x^2+3(1-y)x-(ay+4)=0$$

For x real, D≥0.

$$\Rightarrow 9(1-y)^2+4(a+4y)X(ay+4)\geq 0$$

$$\Rightarrow (9+16a)y^2+\left(4a^2+46\right)y+(9+16a)\geq 0$$

This is a quadratic inequality in y. This inequality holds when coefficient of “y^2 ” is positive and discriminant is less than equal to zero i.e. D’≤0.

$$\Rightarrow (9+16a)>0$$

$$\Rightarrow a>-\frac{9}{16}$$

and

$$D\prime=\left(4a^2+46\right)^2-4(9+16a)^2\le 0$$

$$\Rightarrow \left(2a^2+23\right)^2-(9+16a)^2\le 0$$

$$\Rightarrow \left(2a^2+23+9+16a\right)\left(2a^2+23-9-16a\right)\le 0$$

$$\Rightarrow \left(2a^2+16a+32\right)\left(2a^2-16a+14\right)\le 0$$

$$\Rightarrow \left(a^2+8a+16\right)\left(a^2-8a+7\right)\le 0$$

$$\Rightarrow (a+4)^2\left(a^2-8a+7\right)\le 0$$

$$\Rightarrow \left(a^2-8a+7\right)\le 0$$

$$\Rightarrow \left(a^2-8a+7\right)\le 0$$

The roots of corresponding quadratic equation in a is :

$$\Rightarrow a=1{,}7$$

Hence interval of a satisfying inequality is [1,7]
We have two intervals i.e. a>-9/16 and $\Rightarrow a\in[1{,}7]$ corresponding to two simultaneous conditions. Therefore, values of a is intersection of these two intervals :

$$\Rightarrow a\in[1{,}7]$$

Range of rational function

The set of real values of rational polynomial for real values of x need not be the range of the function. It is because rational function is not defined for zeroes of polynomial in denominator. In previous section, we evaluated values of function for real x. But, domain may not be the real number set, but subset of R, which excludes certain values of x. We need to exclude values of “y”, which corresponds to values of x for which denominator becomes zero. This statement, however, is slightly confusing, because function is not defined for those values of x in the first place. How would we determine values of y corresponding to values of x for which function reduces to indeterminate form involving division by zero. We actually determine limiting values of function at these points and exclude those values of y from the real set of y, which is determined assuming x belonging to R.

There are certain cases in which denominator of the rational function can not become zero. Consider rational functions :

$$f(x) = \frac{2x^2 - x + 1}{x^2 + 1}$$

$$g(x) = \frac{x + 1}{2x^2 - x + 1}$$

$$h(x) = \frac{2x^2 - x + 1}{|x| + 1}$$

The denominators of all these functions can not be zero. Under this condition, domain of the function is real number set R.

Example:

Problem : Find the range of function :

$$f(x) = \frac{x}{1 + x^2}$$

Solution : The denominator of the given rational function can not be zero. Hence, domain of function is real number set R. There is no exclusion point. Rearranging to form a quadratic equation in x, we have :

$$\Rightarrow y+yx^2=x$$

$$\Rightarrow yx^2-x+y=0$$

We should analyze for coefficient of "x^2" in the quadratic equation. For quadratic equation, coefficient of "x^2" can not be zero i.e. y ≠ 0. For real x, y ≠ 0 and D≥0 :

$$D=(-1)^2-4XyXy=1-4y^2\geq 0$$

$$\Rightarrow y^2\geq \frac{1}{4}$$

$$\Rightarrow y\in\left[-\frac{1}{2},\frac{1}{2}\right]$$

What if y=0? Putting this value in the quadratic equation, we have :

$$\Rightarrow 0-x+0=0$$

$$\Rightarrow x=0$$

This is included in the domain. Hence, y=0 is included in the range. The range of the rational function, therefore, remains unaffected :

$$\Rightarrow y\in\left[-\frac{1}{2},\frac{1}{2}\right]$$

Example:

Problem : Find the range of the function :

$$y = f(x) = \frac{x^2 - 5x + 4}{x^2 - 3x + 2}$$

Solution : We see that discrimanants of numerator and denominator polynomials are positive. On factorizing,

$$\Rightarrow y = \frac{x^2 - 5x + 4}{x^2 - 3x + 2} = \frac{(x-1)(x-4)}{(x-1)(x-2)}$$

Clearly, rational function is not defined for x=1 and x=2. Domain of the function is R- {1,2). For the sake of determining range, the limiting values of function for these values of x are obtained by canceling (x-1) from numerator and denominator :

$$\Rightarrow y = \frac{(x-4)}{(x-2)}$$

For x=1, y = 3. For x=2, however, the function value is indeterminate. In totality, we need to exclude y=3 from the interval of real values of y. Now, in order to determine real values of y, we rearrange the given function to form a quadratic equation in x :

$$\Rightarrow yx^2 - 3yx + 2y = x^2 - 5x + 4$$

$$\Rightarrow (y-1)x^2 + (5-3y)x + 2y - 4 = 0$$

We should analyze for coefficient of “x^2 ” in the quadratic equation. For quadratic equation, coefficient of “x^2” can not be zero i.e. y-1 ≠ 0. For real x, y-1 ≠ 0 and D≥0.
For y-1 = 0, y = 1. Putting this value in the quadratic equation,

$$\Rightarrow 0 + (5-3)x + 2 - 4 = 0$$

$$\Rightarrow x = 1$$

We see that x=1 is not part of domain. This is actually the value which reduces denominator to zero. Hence, we should exclude y = 1 from the real values of y. Now for D≥0,

$$D = (5-3y)^2 - 4(y-1)(2y-4) \geq 0$$

$$\Rightarrow 25 + 9y^2 - 30y - 4\left\{2y^2 - 6y + 4\right\} \geq 0$$

$$\Rightarrow 25 + 9y^2 - 30y - 4\left\{2y^2 - 6y + 4\right\} \geq 0$$

The coefficient of

$$y^2$$

is positive. The discriminant is 0. Clearly, following sign rule, f(x) ≥0 for all real values of y. Hence, real values of y are real number set R. However, we need to exclude y = {1,3) as discussed above. Therefore, range of given function is R-{1,3}.

Alternative

Once, exception points are noted, we can evaluate “y” from the reduced form :

$$\Rightarrow y = \frac{(x-4)}{(x-2)}$$

Solving,

$$\Rightarrow x = \frac{2y-4}{(y-1)}$$

Clearly, y#1. But we have seen that y#3 as well. Hence, range of rational function is R-{1,3}.

Graph of rational function

We know that rational function is a composition of two functions in the following form,

$$f(x) = \frac{p(x)}{q(x)}$$

where q(x) ≠ 0. If q(x) = 0, then the ratio has the form “ $x/0$ ”, which is not defined.

For plotting, let us consider a simple rational function given by, $f(x) = 1/x$. This function is known as reciprocal function. It is not defined for x = 0. In order to plot the function, we calculate few initial values as :

$$For \quad x = -1, \quad y = -1$$

$$For \quad x = -2, \quad y = -0.5$$

$$For \quad x = -3, \quad y = -0.33$$

$$For \quad x = 0, \quad \text{y is not defined}$$

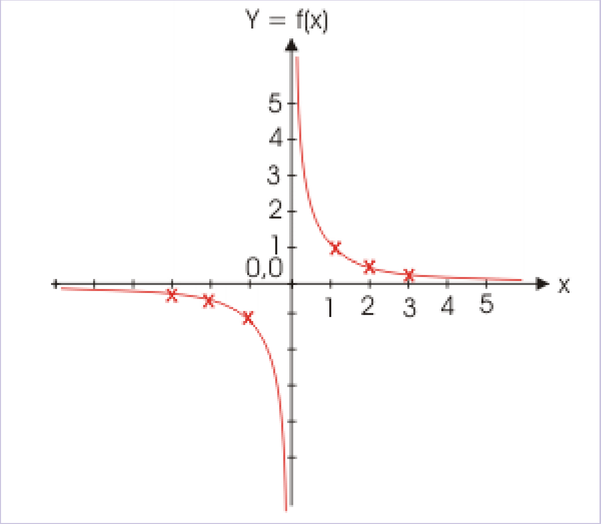
$$For \quad x = 1, \quad y = 1$$

$$For \quad x = 2, \quad y = 0.5$$

$$For \quad x = 3, \quad y = 0.33$$

The graph of the function is shown here :

Rational function



This plot is not defined at x = 0.

This plot is not defined at x = 0. The domain of the given function, therefore, is real numbers, “R” except zero. Also,

$$\Rightarrow x = \frac{1}{y}$$

This means that function value can not be zero. Hence, range of the function is also real numbers, “R” except zero.

$$Domain = R - \{0\}$$

$$Range = R - \{0\}$$

Example:
Problem : Draw the graph of rational function given by :

$$f(x) = \frac{x^2 - 1}{x - 1}$$

Discuss the nature of graph and also determine domain and range of the given function.
Solution : The form of the given function is that of rational function. We observe that the function is not defined for "x = 1" as function has the form " $x/0$ ", which is undefined. The domain of the given function, therefore, is “R” except “1”. It should be noted that while interpreting domain or range we should not cancel out common terms in the numerator and denominator.
For other values of “x”, the value of the function is given by the reduced expression :

$$f(x) = \frac{x^2 - 1}{x - 1} = x + 1$$

Clearly, if the given function were valid for x =1, then y = x+1 = 1 + 1 = 2. Thus, function f(x) can take any real value except “2”. Hence, range of the function is "R" except "2". The domain and range of the given function are :

$$Domain = R - \{1\}$$

$$Range = R - \{2\}$$

In order to plot the function, we calculate few initial values as :

$$For \quad x = -3, \quad y = -2$$

$$For \quad x = -2, \quad y = -1$$

$$For \quad x = -1, \quad y = 0$$

$$For \quad x = 0, \quad y = 1$$

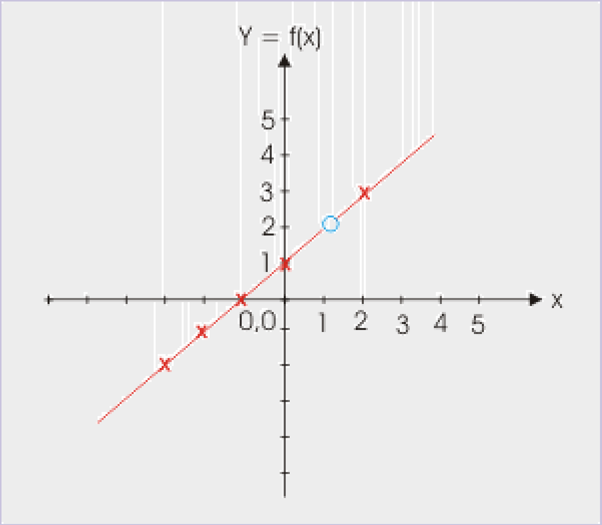
$$For \quad x = 1, \quad \text{y is not defined}$$

$$For \quad x = 2, \quad y = 3$$

$$For \quad x = 3, \quad y = 4$$

The graph of the function is shown here :

Rational function



The plot is not defined at x = 1.

The plot is not defined at x = 1. There is a break at x = 1.

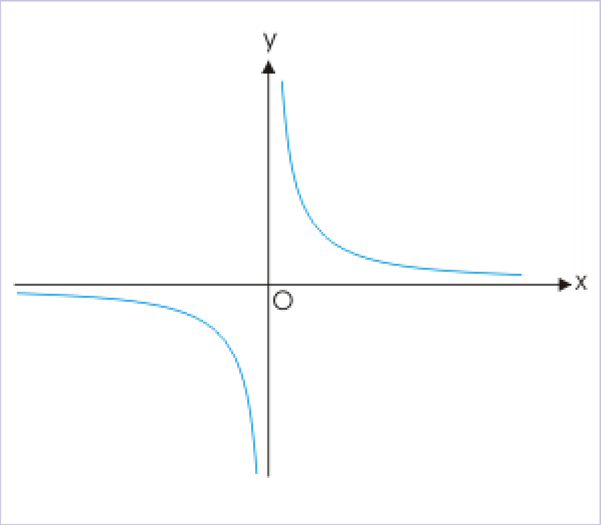
Nature of graph

Here, we consider graphs of rational functions of type :

$$y = \frac{1}{x}, \quad \frac{1}{x^3}, \quad \frac{1}{x^5} \ ,\ldots\ldots\ldots\ldots$$

The nature of graph of these rational function of type $y = \frac{1}{x^n}$, where n is an odd integer such that n≥ 1, is similar to graph of y=1/x as shown in the figure. The graph is that of rectangular hyperbola.

Rational function

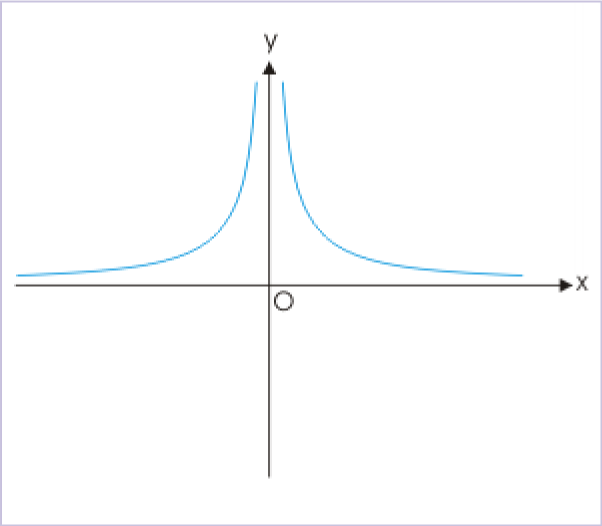


Graph of rational function.

We need to emphasize that the graph generalizes the nature and is helpful to estimate domain and range of functions. We need to graph individual function if required. The nature of graph of function type $y = \frac{1}{x^n}$, where n is an even integer such that n≥2 is shown in the figure below :

$$y = \frac{1}{x^2}, \quad \frac{1}{x^4}, \quad \frac{1}{x^4} \;, \ldots\ldots\ldots\ldots$$

Rational function



Graph of rational function.

Rational inequality

Rational inequality is an inequality involving rational expression. There are four forms of inequality. Corresponding to each of these forms, there are four rational inequality forms. These inequality forms essentially compare a rational expression, f(x), with zero. The four inequalities are :

$$f(x) < 0$$

$$f(x) \leq 0$$

$$f(x) > 0$$

$$f(x) \geq 0$$

We need to note two important aspects of these inequalities. Solution of inequalities, in general, are not discrete values but set of “x” values in the form of interval or union of intervals. Generally, the inequality holds for a continuum of values. Second aspect is about the basic nature of inequality. We know that zero has special significance in real number system. It divides real number system in positive and negative segments. Therefore, solution of these inequalities is about knowing the sign of function values for different intervals in the domain of the function. Corresponding to four inequalities, we need to know intervals in which rational function is (i) negative (ii) non-positive (iii) positive and (iv) non-negative. In the following section, we shall devise a technique to determine sign of rational expression in different internals.

Sign scheme or diagram for rational function

Sign scheme or diagram is representation of sign in different intervals along real number line. This gives a visual idea about the sign of function. Graphically, sign of function changes when graph crosses x-axis. This means that sign of function changes about the zeroes of function i.e. about real roots of a function. However, rational function is ratio of two functions. A change of sign of either numerator or denominator affects sign of rational function.

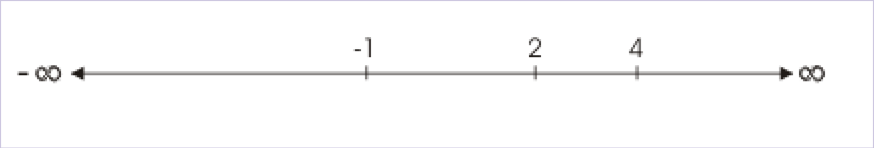
We consider here only integral rational functions such that expressions in numerator and denominator can be decomposed into linear factors. Equating each of the linear factors, we determine points about which either or both of numerator and denominator functions change sign. We should understand that each of the linear factors is a potential source of sign change as the value of x changes in the domain. This means that each of the points so determined plays a critical role in deciding the sign of function. For this reason, we call these points as “critical points”.

Let us consider an example here :

$$f(x)=\frac{x^2-x-2}{x^2-3x-8}=\frac{(x+1)(x-2)}{(x+1)(x-4)}$$

Critical points are -1, 2, -1 and 4. There are two important things to realize here. First, we can not cancel common linear factors as this will result in loosing undefined points and will loose information on sign change. The marking on real number line is as shown here :

Sign scheme/ diagram



Sign scheme/ diagram

Second, the fact that function may change its sign in the domain has an interesting consequence. It can be better understood in terms of function graph, which is essentially a curve. The event of crossing of x-axis by the graph records the event of change of sign. Another change in the sign of graph warrants that curve should cross x-axis again. This corresponds to reversal of sign. It is not possible to change sign of function without crossing x-axis. This means that function will change sign at critical points. Equivalently, we say that sign of function alternates in consecutive sub-intervals. Now, these considerations set up the first two steps of sign diagram :

1: Decompose both numerator and denominator into linear factors. Find critical points by equating linear factors individually to zero.

2: Mark critical points on a real number line. If n be the numbers of critical points, then real number line is divided into (n+1) sub-intervals.

The question however remains that we should know sign of function in at least one interval. We determine the same by testing function value for an intermediate x-value in any of the sub-intervals. Though it is not a rule, we consider a test point in the right most interval, which extends to positive infinity. This helps us to assign signs in the intervals left to it by alternating signs. Sometimes, it may, however, be easier to evaluate function value at x= 0, 1 or -1, provided they are not the critical points. This has the advantage that calculation of function value is easier. Now, these consideration set up the next step of sign diagram :

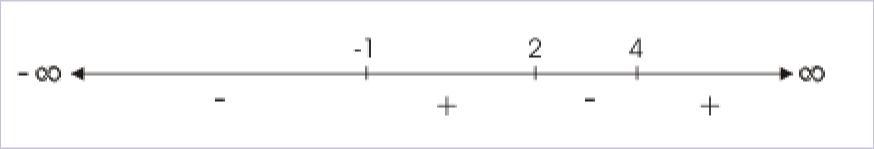
3: Test sign of function in a particular interval. Assign alternate signs in adjacent sub-intervals.

For the example case, let us put x=0,

$$\Rightarrow f(0) = \frac{0-0-2}{0-0-8} = \frac{1}{4} > 0$$

Thus, sign of function in the interval between -1 and 2 is positive. The signs of function alternate in adjacent sub-intervals.

Sign scheme/ diagram



Sign scheme/ diagram

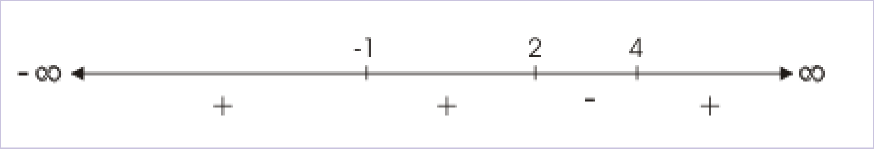
We have noted that sign of each linear factor combines to determine the sign of rational function. This fact is reflected as sign alternates in adjacent sub-intervals. However, we need to consider the effect of case in which a linear factor is repeated. If a linear factor evaluates to a positive number in an interval and is repeated, then there is no effect on the sign of function. If a linear factor evaluates to a negative number in an interval and is repeated even times, then there is no effect on the sign of function. The product of negative sign repeated even times yield a positive sign and as such does not affect the sign of function. However, if a linear factor evaluates to a negative number in an interval and is repeated odd times, then sign of function changes. Product of negative sign repeated odd times yield a negative sign and as such sign of function changes.

We conclude that if a linear factor is repeated even times, then sign of function will not alternate about the critical point corresponding to linear factor in question. On the other hand, if a linear factor is repeated odd times, then sign of function will alternate as before. Now, these consideration set up the next step of sign diagram :

4: If a linear factor is repeated even times, then sign of function will not alternate about the critical point corresponding to linear factor in question.

In the example case, the linear factor (x+1) is repeated even times (count both in numerator and denominator). As such, sign of function will not change about critical point “-1”. Thus, sign diagram drawn as above need to be modified as :

Sign scheme/ diagram



Sign scheme/ diagram

We can verify modification due to repeated linear factors by putting x = -2 in the function :

$$\Rightarrow f(-2)=\frac{(-2)^2-(-2)-2}{(-2)^2-3(-2)-8}=\frac{4+2-1}{4+6-8}=\frac{5}{2}>0$$

We summarize steps for drawing sign scheme/ diagram as :

1: Decompose both numerator and denominator into linear factors. Do not cancel common linear factors. Find critical points by equating linear factors individually to zero.

2: Mark distinct critical points on a real number line. If n be the numbers of distinct critical points, then real number line is divided into (n+1) sub-intervals.

3: Test sign of function in a particular interval. Assign alternate signs in adjacent sub-intervals.

4: If a linear factor is repeated even times, then sign of function will not alternate about the critical point corresponding to linear factor in question.

Solution of rational inequalities using sign scheme or diagram

An important point about interpreting sign diagram is that sign of function relates to non-zero values of function. Note that zero does not have sign. The critical points corresponding to numerator function are zeroes of rational function. As such, the graph of function is continuous at these critical points and these critical points can be included in the sub-interval. On the other hand, the rational function is not defined for critical points corresponding to denominator function (as denominator turns zero). We, therefore, conclude that an interval can include critical points corresponding to numerator function, but not the critical points corresponding to denominator function. In case, there are common critical points between numerator and denominator, then those critical points can not be included in the sub-interval.

We can interpret sign diagram in two ways. Either we determine the solution of a given quadratic inequality or we determine intervals of all four types of inequalities for a given quadratic expression. We shall illustrate these two approaches by working with the example case.

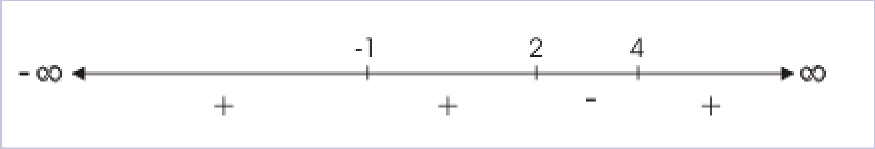
Determining solution of a given quadratic inequality

Let us consider that we are required to solve rational inequality

$$f(x) = \frac{x^2 - x - 2}{x^2 - 3x - 8} \geq 0$$

The sign diagram as drawn earlier for the given rational function is shown here :

Sign scheme/ diagram



Sign scheme/ diagram

We need to interpret signs of different intervals to find the solution of a given rational inequality.

Clearly, solution of given inequality is :

$$x \in (-\infty, 2]U(4,\infty) - \{-1, 4\}$$

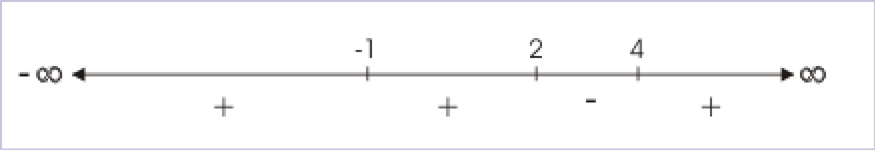
Note that we need to remove -1 and 4 from the solution set as function is not defined for this x – value. However, inequality involved “greater than or equal to” is not strict inequality. It allows equality to zero. As such, we include critical point “2” belonging to numerator function. Further, we can also write the solution set in alternate form as :

$$x \in (-\infty, -1)U(-1,2]U(4,\infty)$$

Determining interval of four quadratic inequalities

Let us take the rational expression of example case and determine intervals of each of four inequalities. The sign diagram as drawn earlier is shown here :

Sign scheme/ diagram



Sign scheme/ diagram

$$f(x) < 0; \quad x \in (2{,}4)$$

$$f(x) \le 0; \quad x \in [2{,}4)$$

$$f(x) > 0; \quad x \in (-\infty, -1)U(-1{,}2)U(4{,}\infty)$$

$$f(x) \ge 0; \quad x \in (-\infty, -1)U(-1{,}2]U(4{,}\infty)$$

Note that critical point “2” belonging to numerator is included for inequalities which allows equality.

Inclusion and exclusion of critical points

Based on the discussion above, we summarize inclusion or exclusion of critical points here :

1: Question of inclusion of critical points arises when inequality involved is not strict.

2: Critical points belonging to numerator are included in solution set.

3: Critical points belonging to denominator are excluded from solution set.

3: Critical points belonging to both numerator and denominator are excluded from solution set.

Solution of rational inequalities using wavy curve method

Wavy curve method is a modified sign diagram method. This method has the advantage that we do not need to test sign of interval as required in earlier case. The steps involved are :

1: Factorize numerator and denominator into linear factors.

2: Make coefficients of x positive in all linear factors. This step may require to change sign of “x” in the linear factor by multiplying inequality with -1. Note that this multiplication will change the inequality sign as well. For example, “less than” will become “greater than” etc.

3: Equate each linear factor to zero and find values of x in each case. The values are called critical points.

4: Identify distinct critical points on real number line. The “n” numbers of distinct critical points divide real number lines in (n+1) sub-intervals.

5: The sign of rational function in the right most interval is positive. Alternate sign in adjoining intervals on the left.

5: If a linear factor is repeated even times, then sign of function will not alternate about the critical point corresponding to linear factor in question.

We need to exclude exception points i.e. critical points of denominator from solution set. Further, it is important to understand that signs of intervals as determined using this method are not the signs of function – rather signs of modified function in which sign of “x” has changed. However, if we are not required to change the sign of “x” i.e. to modify the function, then signs of intervals are also signs of function. We shall though keep this difference in mind, but we shall refer signs of intervals as sign scheme or diagram in this case also.

Example:

Problem : Apply wavy curve method to find the interval of x for the inequality given :

$$\frac{x}{1-x} \geq 0$$

Solution : We change the sign of "x" in the denominator to positive by multiplying both sides of inequality with -1. Note that this changes the inequality sign as well.

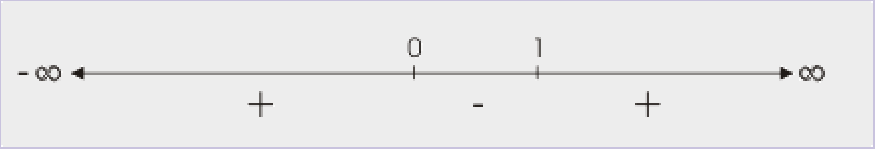
$$\frac{x}{x-1} \leq 0$$

Here, critical points are :

$$x = 0, 1$$

The critical points are marked on the real number line. Starting with positive sign in the right most interval, we denote signs of adjacent intervals by alternating sign.

Sign diagram



Sign of function alternates.

Thus, interval of x as solution of inequality is :

$$\Rightarrow 0 \leq x < 1$$

We do not include "1" as it reduces denominator to zero.

Example:

Problem : Find solution of the rational inequality given by :

$$\frac{3x^2+6x-15}{(2x-1)(x+3)} \geq 1$$

Solution : We first convert the given inequality to standard form f(x) ≥ 0.

$$\Rightarrow \frac{3x^2+6x-15}{(2x-1)(x+3)} - 1 \geq 0$$

$$\Rightarrow \frac{3x^2+6x-15-(2x-1)(x+3)}{(2x-1)(x+3)} \geq 0$$

$$\Rightarrow \frac{3x^2+6x-15-\left(2x^2+5x-3\right)}{(2x-1)(x+3)} \geq 0$$

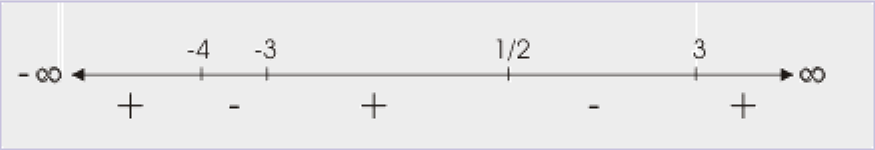
$$\Rightarrow \frac{x^2+x-12}{(2x-1)(x+3)} \geq 0$$

$$\Rightarrow \frac{x^2+4x-3x-12}{(2x-1)(x+3)} \geq 0$$

$$\Rightarrow \frac{(x-3)(x+4)}{(2x-1)(x+3)} \geq 0$$

Critical points are -4, -3, 1/2, 3. Corresponding sign diagram is :

Sign diagram



Sign of function alternates.

The solution of inequality is :

$$\Rightarrow x \in (-\infty, -4] \cup (-3,1/2) \cup [3,\infty)$$

We do not include "-3" and "1" as they reduce denominator to zero.

Example:

Problem : Find solution of :

$$\frac{2}{1+x} + \frac{3}{1-x} < 1$$

Solution : Rearranging, we have :

$$\Rightarrow \frac{2-2x+3+3x}{(1+x)(1-x)} - 1 < 0$$

$$\Rightarrow \frac{5+x-\left(1-x^2\right)}{(1+x)(1-x)} < 0$$

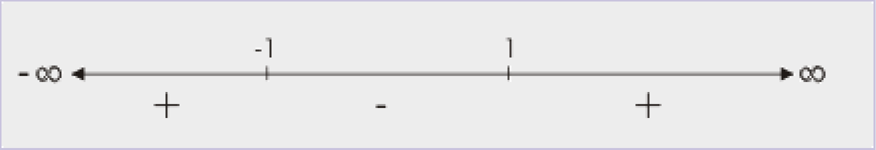
$$\Rightarrow \frac{x^2+x+4}{(1+x)(1-x)} < 0$$

Now, polynomial in the numerator i.e. x^2+x+4 is positive for all real x as D<0 and a>0. Thus, dividing either side of the inequality by this polynomial does not change inequality. Now, we need to change the sign of x in one of the linear factors of the denominator positive in accordance with sign rule. This is required to be done in the factor (1-x). For this, we multiply each side of inequality by -1. This change in sign accompanies change in inequality as well :

$$\Rightarrow \frac{1}{(1+x)(1-x)} > 0$$

Critical points are -1 and 1. Hence, solution of the inequality in x is :

Sign diagram



Sign of function alternates.

$$x \in (-\infty, -1) \cup (1,\infty)$$

Rational inequality with repeated linear factors

We have already discussed rational polynomial with repeated factors. We need to count repeated factors which appear in both numerator and denominator. If the linear factors are repeated even times, then we do not need to change sign about critical point corresponding to repeated linear factor.

Note : While working with rational function having repeated factors, we need to factorize higher order polynomial like cubic polynomial. In such situation, we can employ a short cut. We guess one real root of the cubic polynomial. We may check corresponding equation with values such as 1,2, -1 or -2 etc and see whether cubic expression becomes zero or not for that value. If one of the roots is known, then cubic expression is f(x) = (x-a) g(x), where "a" is the guessed root and g(x) is a quadratic expression. We can then find other two roots anlayzing quadratic expression. For example, $x^3 - 6x^2 + 11x - 6 = (x-1)\left(x^2 - 5x + 6\right) = (x-1)(x-2)(x-3)$

Example:
Problem : Find interval of x satisfying the inequality given by :

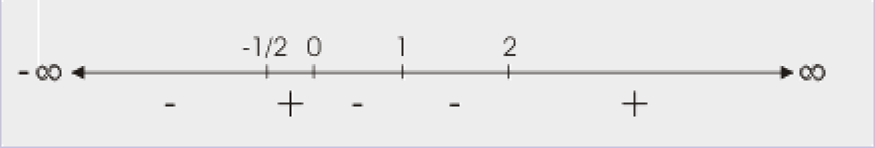
$$\frac{(2x+1)(x-1)}{(x^3-3x^2+2x)} \geq 0$$

Solution : We factorize each of the polynomials in numerator and denominator :

$$\Rightarrow \frac{(2x+1)(x-1)}{(x^3-3x^2+2x)} = \frac{(2x+1)(x-1)}{x(x-1)(x-2)}$$

It is important that we do not cancel common factors or terms. Here, critical points are -1/2,1,0,1 and 2. The critical point "1" is repeated even times. Hence, we do not change sign about "1" while drawing sign scheme.

Sign diagram



Sign of function alternates.

While writing interval, we drop equality sign for critical points, which corresponds to denominator.

$$-1/2 \leq x < 0 \quad \cup \quad 2 < x < \infty$$

$$[-1/2, 0) \quad \cup \quad (2, \infty)$$

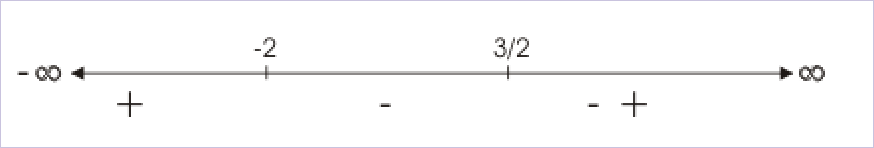
We do not include "-1" and "1" as they reduce denominator to zero.

Polynomial inequality

We can treat polynomial inequality as rational inequality, because a polynomial function is a rational function with denominator as 1. Logically, sign method used for rational function should also hold for polynomial function. Let us consider a simple polynomial inequality,

$f(x) = 2x^2 + x - 1 < 0$. Here, function is product of two linear factors (2x-3)(x+2). Clearly, x=3/2 and x=-2 are the critical points. The sign scheme of the function is shown in the figure :

Sign diagram



Sign of function alternates.

Solution of x satisfying inequality is :

$$x \in \left(-2, \frac{3}{2}\right)$$

It is evident that this method is easier and mechanical in approach.

Radical function

The term radical is name given to square root sign (√). A radical number is n^{th} root of a real number. If y is n^{th} root of x, then :

$$x = y^n$$

Interchangeably, we write :

$$y = x^{\frac{1}{n}} = \sqrt[n]{}(x)$$

If n is even integer, then x can not be negative. For n=2, we drop “n” from the notation and we write,

$$y = \sqrt{x}$$

We extend this concept to function in which number “x” is substituted by any valid expression (algebraic, trigonometric, logaritmic etc). Some examples are :

$$y = \sqrt{x^2 + 3x - 5}$$

$$y = \sqrt{\log_e (x^2 + 3x - 5)}$$

We shall also include study of radical function which is part of rational form like :

$$y = \sqrt{\left(\frac{1}{x_2 + 3x - 5} \right)}$$

$$y = \sqrt{\left\{ \frac{(x-1)(x+3)}{x_2 + 3x - 5} \right\}}$$

Analysis of radical function

Analysis of root function is same as analysis of inequality of function. Because, radical function ultimately results in inequality. We make use of the fact that expression within the radical sign is non-negative. Here, we denote a radical function as :

$$f(x) = \sqrt{g(x)}$$

As the expression under is non-negative,

$$\Rightarrow g(x) \ge 0$$

When radical function is part of a function defined in rational form, the radical function should not be zero. Let us consider a function as :

$$f(x) = \frac{1}{\sqrt{g(x)}}$$

As the radical is denominator of the rational expression, expression under radical sign is positive,

$$g(x) > 0$$

We have already worked with inequalities involving polynomial and rational functions. We shall restrict ourselves to few illustrations here.

Example:
Problem : Find domain of the function :

$$f(x) = \sqrt{\left\{1 - \sqrt{(1 - x^2)}\right\}}$$

Solution : One radical (inner) is contained with another radical (outer). For the outer radical,

$$\Rightarrow 1 - \sqrt{1 - x^2} \geq 0$$

$$\Rightarrow \sqrt{1 - x^2} \leq 1$$

The term on each side of inequality is a positive quantity. Squaring each side does not change inequality,

$$\Rightarrow 1 - x^2 \leq 1$$

$$\Rightarrow x^2 \geq 0$$

This quadratic inequality is true for all real x. Now, for inner radical

$$\Rightarrow 1 - x^2 \geq 0$$

$$\Rightarrow (1 + x)(1 - x) \geq 0$$

We multiply by -1 to change the sign of x in 1-x,

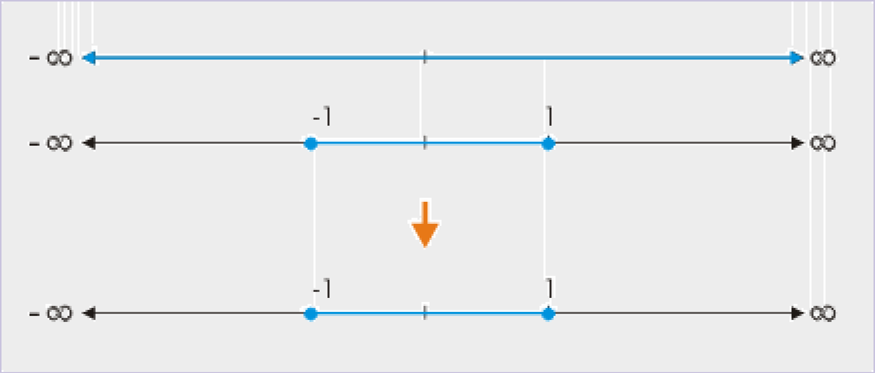
$$\Rightarrow (x+1)(x-1) \le 0$$

Using sign rule :

$$\Rightarrow x \in [-1{,}1]$$

Since, conditions corresponding to two radicals need to be fulfilled simultaneously, the domain of the given function is intersection of outer and inner radicals.

Domain of the function



Intersection of two domains.

$$\text{Domain} = [-1{,}1]$$

Example:

Problem : Find the domain of the function given by :

$$f(x) = \sqrt{(x^{14} - x^{11} + x^6 - x^3 + x^2 + 1)}$$

Solution : Clearly, function is real for values of “x” for which expression within square root is a non negative number. We note that independent variable is raised to positive integers. The nature of each monomial depends on the value of x and nature of power. If x≥1, then monomial evaluates to higher value for higher power. If x lies between 0 and 1, then monomial evaluates to lower value for higher power. Further, a negative x yields negative value when raised to odd power and positive value when raised to even power. We shall use these properties to evaluate the expression for three different intervals of x.

$$\Rightarrow x^{14} - x^{11} + x^6 - x^3 + x^2 + 1 \ge 0$$

We consider different intervals of values of expression for different values of "x", which cover the complete interval of real numbers.

1: $x \ge 1$

In this case, $x^a > x^b, \quad \text{if} \quad a > b$. Evaluating in groups,

$$\left(x^{14} - x^{11}\right) + \left(x^6 - x^3\right) + \left(x^2 + 1\right) > 0$$

2: $0 \le x < 1$

In this case, $x^a < x^b, \quad \text{if} \quad a > b$. Rearranging in groups,

$$x^{14} - \left\{\left(x^{11} - x^6\right) + \left(x^3 - x^2\right)\right\} + 1$$

Here, $\left\{\left(x^{11} - x^6\right) + \left(x^3 - x^2\right)\right\}$ is negative. Hence total expression is positive,

$$\Rightarrow x^{14} - \left\{\left(x^{11} - x^6\right) + \left(x^3 - x^2\right)\right\} + 1 > 0$$

3: $x < 0$

Rearranging in groups,

$$\left(x^{14} - x^{11}\right) + \left(x^6 - x^3\right) + \left(x^2 + 1\right)$$

Here, $x^{14}, x^6 \quad \text{and} \quad x^2$ are positive and $x^{11} \quad \text{and} \quad x^3$ is negative. Hence, total expression is positive,

$$\Rightarrow \left(x^{14} - x^{11}\right) + \left(x^6 - x^3\right) + \left(x^2 + 1\right) > 0$$

We see that expression is positive for all values of "x". Hence, domain of the function is :

$$\text{Domain} = R = (-\infty, \infty)$$

Exercise

Exercise:

Problem: Find solution of the rational inequality given by :

$$\frac{(x+1)(x+5)}{(x-3)} \geq 0$$

Solution:

Hint : Critical points are -5,-1 and 3. We need to exclude end corresponding to x=3 as denominator turns zero for this value.

$$[-5,-1] \quad \cup \quad (3,\infty)$$

Exercise:

Problem: Find solution of the rational inequality given by :

$$\frac{8x^2+16x-51}{(2x-3)(x+4)} > 0$$

Solution:

Hint : Critical points are -4,-3,3/2,5/2.

$$\Rightarrow x \in (-\infty, -4) \cup (-3{,}3/2) \cup (5/2{,}\infty)$$

Exercise:

Problem: Find solution of the rational inequality given by :

$$\frac{x^2+4x+3}{x^3-6x^2+11x-6}>0$$

Solution:

Hint : Factorize denominator as $x^3-6x^2+11x-6=(x-1)(x-2)(x-3)$. Critical points are -3,-1,1,2,3.

$$\Rightarrow x \in (-3,-1) \cup (1{,}2) \cup (3{,}\infty)$$

Exercise:

Problem: Find solution of the rational inequality given by :

$$\frac{(2x+1)(x-1)^2}{(x^3-3x^2+2x)} \geq 0$$

Solution:

Hint : Factorize denominator as

$$\Rightarrow \frac{(2x+1)(x-1)^2}{(x^3-3x^2+2x)} = \frac{(2x+1)(x-1)^2}{x(x-1)(x-2)}$$

Critical points are -1/2,1,1,0,1 and 2. We see that "1" is repeated odd times. Hence, we continue to assign alternating signs in accordance with wavy curve method. The solution of x for the inequality is :

$$-\infty < x \leq -1/2 \quad \cup \quad 0 < x < 1 \quad \cup \quad 2 < x < \infty$$

Function operations

The form of “function as a rule” suggests that we may think of carrying out arithmetic operations like addition, multiplication etc with two functions. If we limit ourselves to real function, then we can attach meaning to equivalent of arithmetic operations with predictable domain intervals. We should, however, clearly understand that function operations with real functions involve new domain for the resulting function. In general, function operation results in contraction of intervals in which new rule formed from algebraic operation is valid.

As pointed out, function operations are defined for a new domain, depending on the type of operations - we carry out. In the nutshell, we may keep following aspects in mind, while describing function operations :

- Function operations are defined for real functions.
- The result of function operation is itself a real function.
- New function resulting from function operation is defined in new interval(s) of real numbers as determined by the nature of operation involved.

Domain of resulting function

The function operations, like addition, involve more than one function. Each function has its domain in which it yields real values. The resulting domain will depend on the way the domain intervals of two or more functions interact. In order to understand the process, let us consider two functions “ y_1 ” and “ y_2 ” as given below.

$$y_1 = \sqrt{(x^2 - 3x + 2)}$$

$$y_2 = \frac{1}{\sqrt{(x^2 - 3x - 4)}}$$

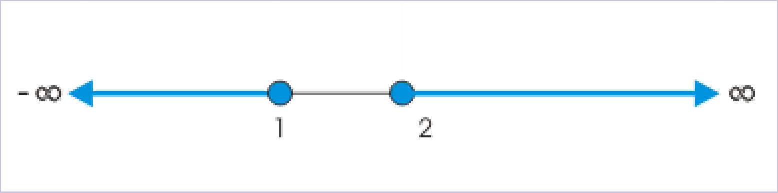
Let " D_1 " and " D_2 " be their respective domains. Now, the expressions in the square roots need to be non-negative. For the first function :

$$x^2 - 3x + 2 \geq 0$$

$$\Rightarrow (x-1)(x-2) \geq 0$$

The sign diagram is shown here. The domain for the function is the intervals in which function value is non-negative.

Domain interval



Domain interval of the first function

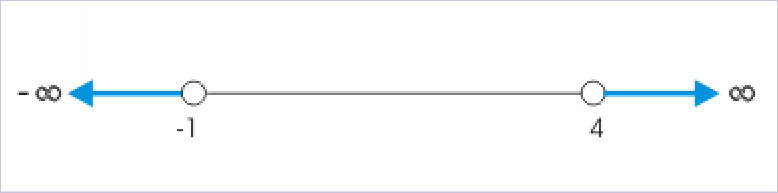
$$\Rightarrow D_1 = -\infty < x \leq 1 \quad \text{or} \quad 2 \leq x < \infty$$

Note that domain, here, includes end points as equality is permitted by the inequality "greater than or equal to" inequality. In the case of second function, square root expression is in the denominator. Thus, we exclude end points corresponding to roots of the equation.

$$x^2 - 3x - 4 > 0$$

$$\Rightarrow (x+1)(x-4) > 0$$

Domain interval



Domain interval of the second function

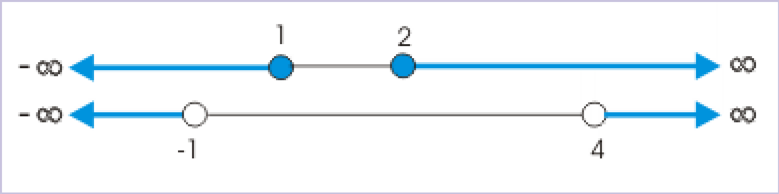
$$\Rightarrow D_2 = -\infty < x < -1 \quad \text{or} \quad 4 < x < \infty$$

Now, let us define addition operation for the two functions as,

$$y = y_1 + y_2$$

The domain, in which this new function is defined, is given by the common interval between two domains obtained for the individual functions. Here, domain for each function is shown together one over other for easy comparison.

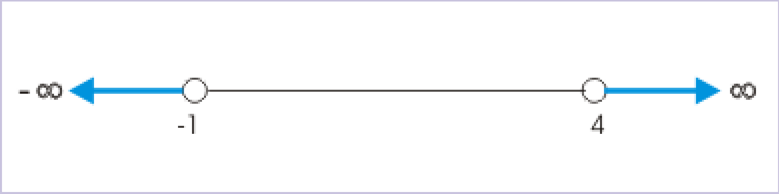
Domain intervals



Domain intervals of two functions

For new function defined by addition operation, values of x should be such that they simultaneously be in the domains of two functions. Consider for example, x = 0.75. This falls in the domain of first function but not in the domain of second function. It is, therefore, clear that domain of new function is intersection of the domains of individual functions. The resulting domain of the function resulting from addition is shown in the figure.

Domain interval



Domain interval of the sum of two functions

$$\Rightarrow D = D_1 \cap D_2$$

This illustration shows how domains interact to form a new domain for the new function when two functions are added together.

Function operations

We, now, define valid operations for real functions in the light of discussion about the domain in the previous section. For different function operations, let us consider two real functions “f” and “g” with domains “ D_1 ” and “ D_2 ” respectively. Clearly, these domains are real number set R or subsets of R :

$$D_{1,}D_2 \in R$$

Addition

The addition of two real functions is denoted as “f+g”. It is defined as :

$$f + g : D_1 \cap D_2 \rightarrow R \quad \text{such that :}$$

$$(f + g)(x) = f(x) + g(x) \quad \text{for all} \quad x \in D_1 \cap D_2$$

Subtraction

The subtraction of two real functions is denoted as “f-g”. It is defined as :

$$f - g : D_1 \cap D_2 \rightarrow R \quad \text{such that :}$$

$$(f - g)(x) = f(x) - g(x) \quad \text{for all} \quad x \in D_1 \cap D_2$$

Scalar Multiplication

Scalar, here, means a real number constant, say “a”. The scalar multiplication of a real function with a constant is denoted as “af”. It is defined as :

$$af : D_1 \rightarrow R \quad \text{such that :}$$

$$(af)(x) = af(x) \quad \text{for all} \quad x \in D_1$$

Multiplication

The product of two real functions is denoted as “fg”. It is defined as :

$$fg : D_1 \cap D_2 \rightarrow R \quad \text{such that :}$$

$$(fg)(x) = f(x)g(x) \quad \text{for all} \quad x \in D_1 \cap D_2$$

Quotient

The quotient of two real functions is denoted as “f/g”. It involves rational form as “f(x)/g(x)”, which is defined for g(x) ≠ 0. We need to exclude value of “x” for which g(x) is zero. Hence, it is defined as :

$$\frac{f}{g} : D_1 \cap D_2 - \left\{x \middle| g(x) \neq 0\right\} \rightarrow R \quad \text{suchthat :}$$

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} \quad \text{for all} \quad x \in D_1 \cap D_2 - \left\{x \middle| g(x) \neq 0\right\}$$

Example

Example:

Problem : Let two functions be defined as :

$$f(x) = \sqrt{x}$$

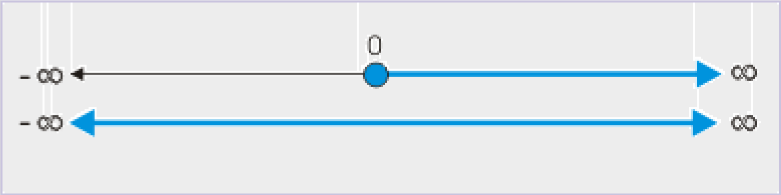
$$g(x) = x^2 - 5x + 6$$

Find domains of “fg” and “f/g”. Also define functions fg(x) and (f/g)(x). **Solution :** The function f(x) is defined for all non negative real number. Hence, its domain is :

$$\Rightarrow x \geq 0$$

The function, g(x), - being a real quadratic polynomial - is real for all real values of “x”. Hence, its domain is “R”. Domains of two functions are shown in the figure.

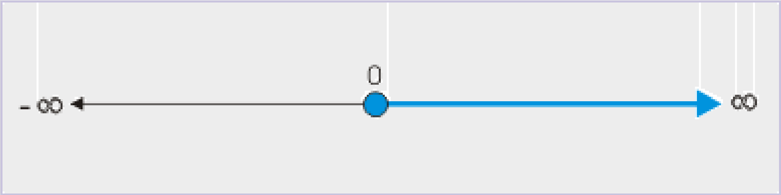
Domain intervals



Domain interval of two functions

The domain of fg(x) is intersection of two intervals, which is non-negative interval as shown in the figure :

Domain interval



Domain interval of product

$$D = D_1 \cap D_2 = [0, \infty] \cap R$$

Here, we recall that intersection of a set with subset is equal to subset :

$$D = [0, \infty]$$

This is the domain of product function “fg(x)”. The domain of quotient function “f/g(x)” excludes values of “x” for which “g(x)” is zero. In other words, we exclude roots of “g(x)” from domain. Now,

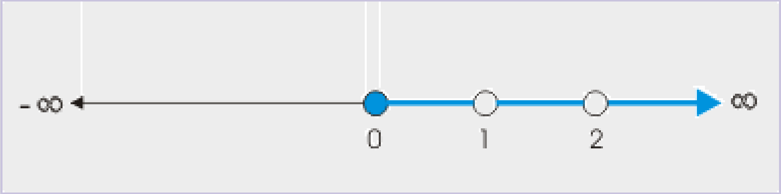
$$g(x) = x^2 - 5x + 6 = 0$$

$$\Rightarrow g(x) = (x-2)(x-3) = 0$$

$$\Rightarrow x = 2{,}3$$

Hence, domain of “f/g(x)” is :

Domain interval



Domain interval of product

$$\Rightarrow D_1 \cap D_2 - \{x | g(x) \neq 0\} = [0{,}\infty] - \{2{,}3\}$$

Now the product function, in rule form, is given as :

$$fg(x) = \sqrt{x}\left(x^2 - 5x + 6\right); \quad x \geq 0$$

Similarly, quotient function, in rule form, is given as :

$$fg(x) = \frac{\sqrt{x}}{x^2 - 5x + 6}; \quad x \geq 0, x \neq 2, x \neq 3$$

Example:

Problem : Find domain of the function :

$$f(x) = 2\sqrt{(x-1)} + \sqrt{(1-x)} + \sqrt{(x_2 + x + 1)}$$

Solution : Given function can be considered to be addition of three separate function. We know that scalar multiplication of a function does not change domain. As such, domain of $2\sqrt{(x-1)}$ is same as that of $\sqrt{(x-1)}$. For $\sqrt{(x-1)}$ and $\sqrt{(1-x)}$,

$$x - 1 \geq 0 \quad \Rightarrow x \geq 1$$

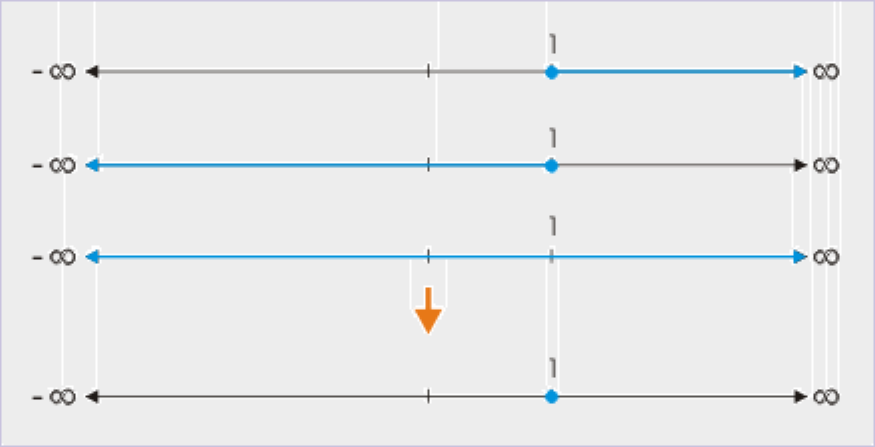
$$1 - x \geq 0 \quad \Rightarrow x \leq 1$$

Now, we use sign rule for third function :

$$x^2 + x + 1 \geq 0$$

Here, coefficient of “x^2 ” is positive and D is negative. Hence, function is positive for all real x. This means f(x)>0. This, in turn, means f(x)≥0. The domain of third function is R. Domain of given function is intersection of three domains. From figure, it is clear that only x=1 is common to three domains. Therefore,

Domain of function



Intersection of three domains.

$$\text{Domain} = \{1\}$$

Exercise

Exercise:

Problem: **Problem :** Find the domain of the function given by :

$$f(x) = \frac{x}{\sqrt{(x^2 - 5x + 6)}}$$

Solution:

Solution :

The function is in rational form. We can treat numerator and denominator functions separately as f(x) and g(x). The numerator is valid for all real values of “x”. Hence, its domain is “R”.

$$D_1 = R$$

For determining domain of g(x), we are required to find the value of “x” for which square root in the denominator is real and not equal to zero. Thus, we need to evaluate square root expression for positive number. It means that :

$$x^2 - 5x + 6 > 0 \quad \Rightarrow (x-2)(x-3) > 0$$

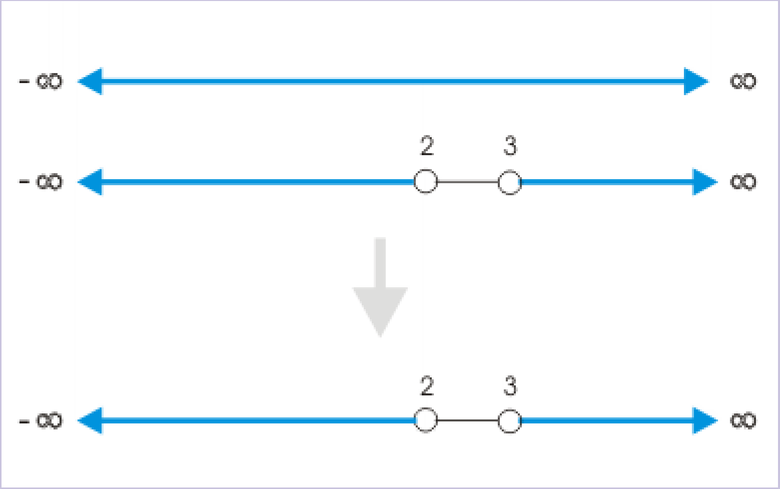
The roots of the corresponding quadratic equation is 2,3. Further, coefficient of "x^2 " term is a positive number (1>0) . Therefore, intervals on the sides are positive for the quadratic expression. The valid interval satisfying the inequality is :

$$x < 2 \quad \text{or} \quad x > 3$$

$$D_2 = (-\infty, 2) \quad \cup \quad (3,\infty)$$

Now, given function is quotient of two functions. Hence, domain of the given function is intersection of two domain excluding interval that renders denominator zero. However, we have already taken into account of this condition, while determining domain of the function in the denominator. Hence,

Domain



Domain of the function is equal to intersection of two domains.

$$\text{Domain} = D_1 \quad \cup \quad D_2 = R \quad \cap \quad \{(-\infty, 2) \quad \cup \quad (3,\infty)\}$$

$$\Rightarrow \text{Domain} = \{(-\infty, 2) \quad \cup \quad (3,\infty)\}$$

Modulus function

The modulus function returns positive value of a variable or an expression. For this reason, this function is also referred as absolute value function. In reference to modulus of an independent variable, the function results in a non-negative value of the variable, irrespective of whether independent variable is positive or negative. The intent of this expression for different values of “x” is expressed as :

`| x, x≥0 |x| = | | -x, x<0`

When value of "x" is a non-negative number, then function is "x"; otherwise "-x". The negative sign in the second interval ensures that the function return a positive value when variable is negative.

There are different interpretations of modulus, depending on situations. In particular, modulus of a real variable and modulus of an expression needs to be interpreted in proper context. Modulus of real variable is equivalent to modulus of real number as variable can take real values only. On the other hand modulus of an expression needs to be treated differently. An expression in a variable is a function. The fact that modulus modifies function values changes function properties. Besides, modulus function can also be interpreted to represent distance of a point with respect to a reference point.

In order to investigate the nature of modulus function, we investigate its plot with respect to independent variable. Here, we calculate few initial values to draw the plot as :

$$For \quad x=-2, \quad y=|x|=-x=-(-2)=2$$

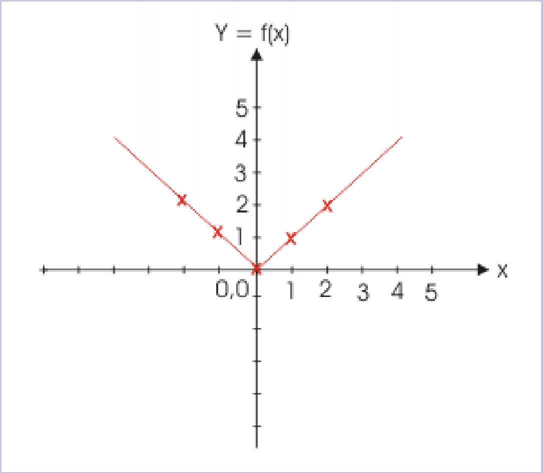
$$For \quad x=-1, \quad y=|x|=-x=-(-1)=1$$

$$For \quad x=0, \quad y=|x|=x=0$$

$$For \quad x=1, \quad y=|x|=x=1$$

$$For \quad x=2, \quad y=|x|=x=2$$

The graph of the function is shown here :
Modulus function



The domain of the function is R.

The graph of modulus function is continuous having a corner at x=0. It means that graph is differentiable at all points except at x=0 (we can not draw a tangent at a corner). Since graph is symmetric about y-axis, modulus function is even function. We also see that there are pair of x values for non-zero values of y. This, in turn, means that image and pre-images are not uniquely related. As such, modulus function is not invertible.

It is clear from the graph that the domain of modulus function is "R". However, the function values are only positive values, including zero. Hence, range of modulus function is upper half of the real number set, including zero.

$$Domain = R$$

$$Range = [0,\infty)$$

In the nutshell, we see that modulus represents a non-negative value. Going by this interpretation, modulus of a variable can also be thought as the square root of the square of the variable. We must emphasize that an unsigned square root is a non-negative value same as modulus value. Hence,

$$|x| = \sqrt{x^2}$$

Modulus function is a non – negative value. There are some other such non-negative expressions. Here, we enumerate them for ready reference :

- Modulus of an expression or variable : |x|
- Even powers of an expression or variable : $x^{2n}, \quad \text{where} \quad n \in Z$
- Even root of an expression or variable : $x^{\frac{1}{2n}}, \quad \text{where} \quad n \in Z$
- y = 1 – sinx; y = 1 – cosx; (as sinx ≤1 and cosx ≤1)

Modulus and equality

Modulus of a variable or an expression is referred frequently in different mathematical contexts. Modulus is also widely used to denote intervals. The representation of interval, in terms of modulus, has the advantage of compactness. We, however, need to be careful in interpreting modulus. Here, we present certain important interpretations of expressions and equations, which involve modulus.

For the sake of understanding, we consider a non-negative number "2" equated to modulus of independent variable "x" like :

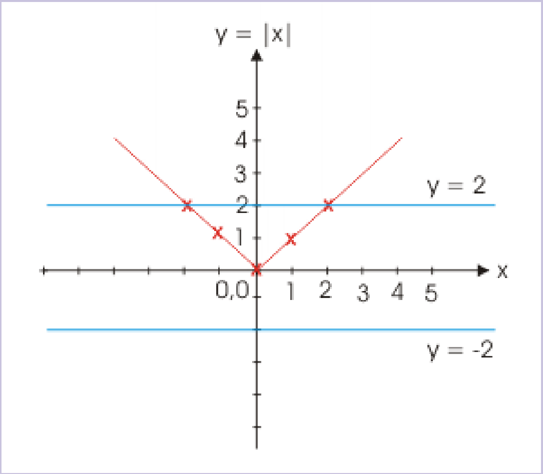
$$|x| = 2$$

Then, the values of "x" satisfying this equation is :

$$\Rightarrow x = \pm 2$$

It is intuitive to note that values of "x" satisfying above equation is actually the intersection of modulus function "y=[x]" and "y=2" plots as shown in the figure.

Modulus function



The values of "x" satisfying modulus equation .

Further, it is easy to realize that equating a modulus function to a negative number is meaningless. The modulus expression "|x|" always evaluates to a non-negative number for all real values of "x". Observe in the figure above that line "y=-2" does not intersect modulus plot at all.

We express these results in general form, using an expression f(x) in place of "x" as :

$$|\mathrm{f}(\mathrm{x})| = a; \quad a > 0 \quad \Rightarrow \mathrm{f}(\mathrm{x}) = \pm a$$

$$|\mathrm{f}(\mathrm{x})| = a; \quad a = 0 \quad \Rightarrow \mathrm{f}(\mathrm{x}) = 0$$

$$|\mathrm{f}(\mathrm{x})| = a; \quad a < 0 \quad \Rightarrow \text{There is no solution of this equality}$$

Modulus as distance

The modulus of an expression “x-a” is interpreted to represent “distance” between “x” and “a” on the real number line. For example :

$$|x-2| = 5$$

This means that the variable “x” is at a distance “5” from “2”. We see here that the values of “x” satisfying this equation is :

$$\Rightarrow x - 2 = \pm 5$$

Either

$$\Rightarrow x = 2 + 5 = 7$$

or,

$$\Rightarrow x = 2 - 5 = -3$$

The “x = 7” is indeed at a distance “5” from “2” and “x=-3” is indeed at a distance “5” from “2”. Similarly, modulus |x|= 3 represents distance on either side of origin.

Modulus and inequality

Interpretation of inequality involving modulus depends on the nature of number being compared with modulus.

Case 1 : a>0

The values of x that satisfy "less than (<)" inequality lies between intervals defined between -a and a, excluding end points of the interval. On the other hand, values of x that satisfy "greater than (>)" inequality lies in two disjointed intervals.

$$|x| < a; \quad a > 0 \quad \Rightarrow -a < x < a$$

$$|x| > a; \quad a > 0 \quad \Rightarrow x < -a \quad \text{or} \quad x > a \quad \Rightarrow x \in (-\infty,\text{a}) \cup (\text{a},\infty)$$

Important aspect of these inequalities is that they can be used to express intervals in compact form. For example, range of cosecant trigonometric function is $x \in (-\infty,\text{-}1] \cup [1,\infty\}$. Equivalently, we can write this interval as |x|≥1.

We extend these results to an expression as :

$$|\text{f(x)}| < a; \quad a > 0 \quad \Rightarrow -a < \text{f(x)} < a$$

For inequality involving greater than comparison with a positive number represents union of two separate intervals.

$$|\text{f(x)}| > a; \quad a > 0 \quad \Rightarrow \text{f(x)} < -a \quad \text{or} \quad \text{f(x)} > a$$

Case 2 : a<0

Modulus can not be equated to negative number as modulus always evaluates to non-negative number. Clearly, modulus of an expression or variable can not be less than a negative number. However, modulus function is always greater than negative number. Hence, we conclude that :

$$|x| < a; \quad a < 0 \quad \Rightarrow \quad \text{There is no solution of this inequality}$$

$$|x| > a; \quad a < 0 \quad \Rightarrow \quad \text{This inequality is valid for all real values of x}$$

Extending these results to expression, we have :

$$|\text{f(x)}| < a; \quad a < 0 \quad \Rightarrow \quad \text{There is no solution of this inequality}$$

$$|\text{f(x)}| > a; \quad a < 0 \quad \Rightarrow \quad \text{This inequality is valid for all real values of f(x)}$$

Example:

Problem : Find the domain of the function given by :

$$f(x) = \frac{x}{\sqrt{(|x| - x)}}$$

Solution : The function is in rational form. The domain of the function in the numerator is "R". We are, now, required to find the value of “x” for which denominator is real and not equal to zero. Now, expression within square root is a non-negative. However, as the function is in denominator, it should not evaluate to zero either. It means that the expression within square root is positive :

$$\Rightarrow |x| - x > 0$$

$$\Rightarrow |x| > x$$

If “x” is a non-negative number, then by definition, "|x| = x". This result, however, is contradictory to the inequality given above. Hence, “x” can not be non-negative. When "x" is a negative number, then the inequality holds as modulus of a real variable is always greater than negative number. It means that "x" is a negative number :

$$\Rightarrow x < 0$$

Thus, domain of the given function is equal to intersection of "R" and interval "x<0". It is given by :

$$\text{Domain} = (-\infty, 0)$$

Additional properties of modulus function

Here, we enumerate some more properties of modulus of a real variable i.e. modulus of a real number.

Let x,y and z be real variables. Then :

$$|-x| = |x|$$

$$|x - y| = 0 \quad \Leftrightarrow x = y$$

$$|x + y| \leq |x| + |y|$$

$$|x - y| \geq ||x| - |y||$$

$$|xy| = |x| \text{ X } |y|$$

$$\left|\frac{x}{y}\right| = \frac{|x|}{|y|}; \quad \left|y\right| \neq 0$$

Modulus |x-y| represents distance of x from y. Also, we know that sum of two sides of a triangle is greater than third side. Combining these two facts, we write a general property for modulus involving real numbers as :

$$|x-y| < |x-z| + |z-y|$$

Square function

There is striking similarity between modulus and square function. Both functions evaluate to non-negative values.

$$y = |x|; \quad y \geq 0$$

$$y = x^2; \quad y \geq 0$$

Their plots are similar. Besides, they behave almost alike to equalities and inequalities. We shall not discuss each of the cases as done for the modulus function, but with a specific number (4 or -4). We shall enumerate each of the possibilities, which can be easily understood in the background of discussion for modulus function.

1: Equality

$$x^2 = 4 \quad \Rightarrow x = \pm 2$$

$$x^2 = -4 \quad \Rightarrow \quad \text{No solution}$$

2: Inequality with non-negative number

A. Less than or less than equal to

$$x^2 < 4 \quad \Rightarrow -2 < x < 2$$

B. Greater than or greater than equal to

$$x^2 > 4 \quad \Rightarrow x < -2 \quad \text{or} \quad x > 2 \quad \Rightarrow \left(-\infty, -2\right) \quad \cup \quad \left(2,\infty\right)$$

3: Inequality with negative number

A. Less than or less than equal to

$$x^2 < -4 \quad \Rightarrow \quad \text{No solution}$$

B. Greater than or greater than equal to

$$x^2 > -4 \quad \Rightarrow \quad \text{Always true}$$

Acknowledgment

Author wishes to thank Mr. Ritesh Shah for making suggestion to remove error in the module.

Greatest and least integer functions

In this module, we shall study a family of functions which return integers based on certain rule, corresponding to a real number. Greatest integer function (floor), least integer function (ceiling) and nearest integer function form part of this family.

Greatest integer function (Floor function)

Greatest integer function returns the greatest integer less than or equal to a real number. In other words, we can say that greatest integer function rounds “down” any number to the nearest integer. This function is also known by the names of “floor” or “step” function. The greatest integer function (GIF) is denoted by the symbol “[x]” .

Interpretation of Greatest integer function is straight forward for positive number. Consider the values “0.23” and “1.7”. The greatest integers for two numbers are “0” and “1”. Now, consider a negative number “-0.54” and “-2.34”. The greatest integers less than these negative numbers are “-1” and “-3” respectively.

We can observe here that greater integer function is actually a function that returns the integral part of a positive real number. This interpretation is clear for positive number. Interpretation for negative numbers needs some explanation. We interpret these values in the context of the fact that every real number can be decomposed to have two parts (i) integral and (ii) fractional part. From this point of view, the negative number can be thought as :

$$\text{-0.54 (real number)} = \text{-1 (integral part)} + \text{0.36 (fraction part)}$$

$$\text{-2.34 (real number)} = \text{-3 (integral part)} + \text{0.66 (fraction part)}$$

We may be tempted to disagree (why not -2 + -0.34 = -2.34?). But, we should know that this is how greatest integer function (GIF) treats a negative number. It returns "-3" for "-2.34" - not "-2". Subsequently, we shall define a function called fraction part function (FPF) that returns fraction part of real number. We shall find that the function exactly returns the same fraction for negative number as has been worked out. The fraction part function (FPF) returns a fraction, which is always positive. It is denoted as {x}. Because of these aspects of GIF and FPF, we can understand the reason why negative number is treated the way it has been presented above. In terms of integral and fraction parts, we write a real number "x" as :

$$x = [x] + \{x\}$$

In the nutshell, we can use any of the following interpretations of greatest integer function :

- [x] = Greatest integer less than equal to “x”
- [x] = Greatest integer not greater than “x”
- [x] = Integral part of “x”

The value of "[x]" is an integer (n) such that :

$$f(x) = [x] = n; \quad \text{if} \quad n \le x < n+1 \quad n \in Z$$

Working rules for evaluating greatest integer function are two step process :

1. If “x” is an integer, then [x] = x.
2. If “x” is not an integer, then [x] evaluates to greatest integer less than “x”.

Graph of greatest integer function

Few initial function values are :

$$For \quad -2 \le x < -1, \quad f(x) = [x] = -2$$

$$For \quad -1 \le x < 0, \quad f(x) = [x] = -1$$

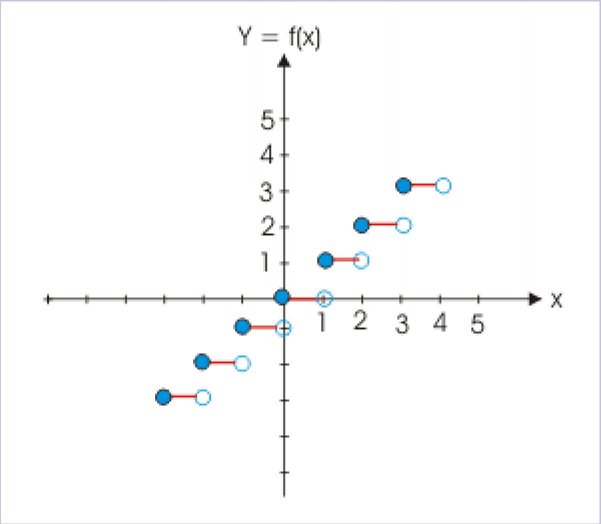
$$For \quad 0 \le x < 1, \quad f(x) = [x] = 0$$

$$For \quad 1 \le x < 2, \quad f(x) = [x] = 1$$

$$For \quad 2 \le x < 3, \quad f(x) = [x] = 2$$

The graph of the function is shown here :

Greatest integer function



The domain of the function is R.

This function is known as step function as values of function steps by "1" as we switch values of “x” from one interval to another. We see that there is no restriction on values of "x" and as such its domain has the interval equal to that of real numbers. On the other hand, the step function or greatest integer function evaluates only to integer values. It means that the range of the function is set of integers, denoted by "Z". Hence,

$$Domain = R$$

$$Range = Z$$

GIF is not a periodic function. Though function is defined for all real x, but graph is not continous. It breaks at integral values of x.

Example:

Problem : Find domain of function given by :

$$f(x) = \frac{1}{\sqrt{\pi - [x]}}$$

Solution : The denominator of function is positive. This means :

$$\Rightarrow \pi - [x] > 0$$

$$\Rightarrow [x] < \pi$$

The value of π is 3.14. Here, [x] returns integral value. Clearly, it can assume a maximum value of 3. But, GIF returns integer value "n" for x<n+1. The inequality, therefore, has solution given by :

$$\Rightarrow x < 4$$

$$\text{Domain} = (-\infty, 4)$$

Important properties

Certain properties of greatest integer function are presented here :

1: If and only if "x"is an integer, then :

$$[x] = x$$

2: If and only if at least either "x" or "y" is an integer, then :

$$[x + y] = [x] + [y]$$

For example, let x = -2.27 and y = 0.63. Then,

$$\Rightarrow [x + y] = [-2.27 + 0.63] = [-1.64] = -2$$

$$\Rightarrow [x] + [y] = [-2.27] + [0.63] = -3 + 0 = -3$$

However, if one of two numbers is integer like x = -2 and y = 0.63, then the proposed identity as above is true.

4: If “x” belongs to integer set, then :

$$[x] + [-x] = 0; \quad x \in Z$$

For example, let x = 2.Then

$$\Rightarrow [2] + [-2] = 2 - 2 = 0$$

We can use this identity to test whether “x” is an integer or not?

3: If “x” does not belong to integer set, then :

$$[x] + [-x] = -1; \quad x \notin Z$$

For example, let x = 2.7.Then

$$\Rightarrow [2.7] + [-2.7] = 2 - 3 = -1$$

Example:

Problem : Find domain of the function :

$$\mathrm{f(x)} = \frac{1}{[x-2]}$$

Solution : Given function is in rational form having GIF as its denominator. The denominator should not evaluate to zero for real values of x. The domain of GIF is real number set R. But, we know that GIF evaluates to zero in an interval which is spread over unit value. In order to know this interval, we determine interval of x for which [x-2] is zero.

$$\Rightarrow [x-2] = 0$$

We can write this function as :

$$\Rightarrow [x + (-2)] = 0$$

Using property [x+y] = [x] + [y], provided one of x and y is an integer. This is the case here,

$$\Rightarrow [x+(-2)] = [x] + [-2] = [x] - 2 = 0$$

$$[x] = 2 \quad \Rightarrow 2 \le x < 3 \quad \Rightarrow x \in [2,3)$$

Hence, domain of given function is :

$$\text{Domain} = R - [2,3)$$

Fraction part function

We define a fraction part function (FPF) denoted by “{x}” as :

$$\{x\} = x - [x]$$

This function returns fraction part of the number, when “x” is not an integer. This exception of non-integral “x” is important. Zero is not a fraction. For integer "x", the function evaluates to zero :

$$\Rightarrow \{5\} = 5 - [5] = 5 - 5 = 0$$

$$\Rightarrow \{-5\} = -5 - [-5] = -5 + 5 = 0$$

Though zero is not a fraction, but FPF evaluates to zero for integral values. We should keep this exception in mind, while working with FPF. Let us, now, work out with numbers that we earlier used for evaluating greatest integer function :

$$\Rightarrow \{0.23\} = 0.23 - [0.23] = 0.23 - 0 = 0.23$$

$$\Rightarrow \{1.7\} = 1.7 - [1.7] = 1.7 - 1 = 0.7$$

$$\Rightarrow \{-0.54\} = -0.54 - [-0.54] = -.54 - (-1) = -0.54 + 1.0 = 0.36$$

$$\Rightarrow \{-2.34\} = -2.34 - [-2.34] = -2.34 - (-3) = -2.34 + 3.0 = 0.66$$

We can see that interpretation of fraction for the negative number is consistent with what has been explained earlier.

Graph of {x}

Few function expressions in different intervals are :

$$\text{For}\quad -2 \leq x < -1, \quad f(x) = \{x\} = x - [x] = x - (-2) = x + 2$$

$$\text{For}\quad -1 \leq x < 0, \quad f(x) = \{x\} = x - [x] = x - (-1) = x + 1$$

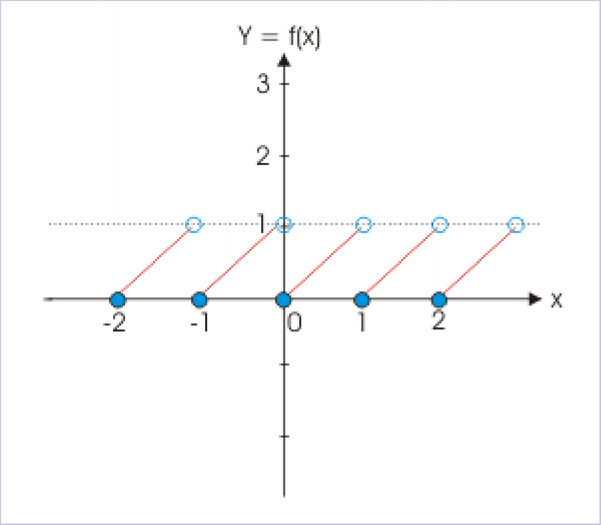
$$\text{For}\quad 0 \leq x < 1, \quad f(x) = \{x\} = x - [x] = x - 0 = x$$

$$\text{For}\quad 1 \leq x < 2, \quad f(x) = \{x\} = x - [x] = x - 1$$

$$\text{For}\quad 2 \leq x < 3, \quad f(x) = \{x\} = x - [x] = x - 2$$

The graph of the function is shown here :

Graph of {x} function



The domain of the function is R.

We see that there is no restriction on values of x and as such its domain has the interval equal to that of real numbers. The fractional part function can only evaluate to non-negative values between 0≤y<1. Hence,

$$\text{Domain} = R$$

$$\text{Range} = 0 \le y < 1$$

FPF is a periodic function. The values are repeated with a period of 1. Further, function is defined for all real x, but graph is not continous. It breaks at integral values of x.

Least integer function

We have seen that greatest integer function represents the integer, which can be considered to be the floor integral value of a real number. Correspondingly, we define a ceiling function called “least integer function (LIF)”, which returns the least integer greater than or equal to the number (x). We denote least integer function as “[x)” or "(x)". Some authors reserve "(x)" for near integer function. It is not important as we can always specify what we mean by qualifying the symbol explicitly. We interpret LIF as :

- [x) = least integer greater than or equal to the number x
- [x) = least integer not less than or equal to the number x

Clearly, least integer function returns a value, which is the integral “ceiling” of the number. For this reason, least integer function is also known as “ceiling” function. Working rules for finding least integer function are :

1. If “x” is an integer, then [x) = x.
2. If “x” is not an integer, then [x) evaluates to least integer greater than “x”.

The value of f(x) is an integer (n) such that :

$$f(x) = n; \quad \text{if} \quad n-1 < x \le n \quad n \in Z$$

Graph of least integer function

Few initial values of the functions are :

$$For \quad -3 < x \le -2, \quad f(x) = [x) = -2$$

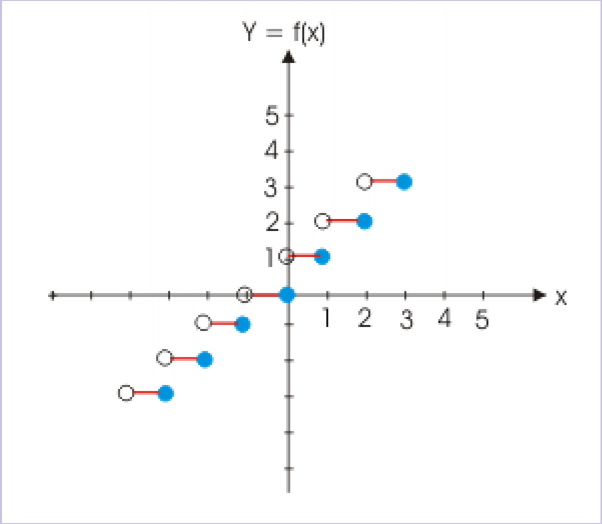
$$For \quad -2 < x \le -1, \quad f(x) = [x) = -1$$

$$For \quad -1 < x \le 0, \quad f(x) = [x) = 0$$

$$For \quad 0 < x \le 1, \quad f(x) = [x) = 1$$

$$For \quad 1 < x \le 2, \quad f(x) = [x] = 2$$

Graph of least integer function



The domain of the function is R.

We see that there is no restriction on values of “x” and as such its domain has the interval equal to that of real numbers. On the other hand, the least integer function evaluates only to integer values. It means that the range of the function is set of integers, denoted by "Z". Hence,

$$\text{Domain} = R$$

$$\text{Range} = Z$$

GIF is not a periodic function. Though function is defined for all real x, but graph is not continous. It breaks at integral values of x.

Important properties

Certain properties of least integer function are presented here :

1: If and only if “x”is an integer, then :

$$[x) = x$$

2: If and only if at least either “x” or “y” is an integer, then :

$$[x + y) = [x) + [y)$$

For example, let x = 2.27 and y = 0.63. Then,

$$\Rightarrow [x + y) = [2.27 + 0.63) = [2.9) = 3$$

$$\Rightarrow [x) + [y) = [2.27) + [0.63) = 3 + 1 = 4$$

However, if one of two numbers is integer like x = 2 and y = 0.63, then the proposed identity as above is true.

4: If “x” belongs to integer set, then :

$$[x) + [-x) = 0; \quad x \in Z$$

For example, let x = 2.Then

$$\Rightarrow [2) + [-2) = 2 - 2 = 0$$

We can use this identity to test whether “x” is an integer or not?

3: If “x” does not belong to integer set, then :

$$[x) + [-x) = +1; \quad x \notin Z$$

For example, let x = 2.7.Then

$$\Rightarrow [2.7) + [-2.7) = 3 - 2 = +1$$

Nearest integer function

Nearest integer function, as the name suggests, returns the nearest integer. It is denoted by the symbol, "(x)".

The value of "(x)" is an integer "n" such that :

$$f(x) = (x) = n; \quad \text{if} \quad n \le x \le n + 1/2, \quad n \in Z$$

$$f(x) = n + 1; \quad \text{if} \quad n + 1/2 < x \le n + 1, \quad n \in Z$$

Examples :

$$(2.3) = 2, \quad (2.6) = 3$$

$$(-2.3) = -2, \quad (-2.6) = -3$$

Exercise

Exercise:

Problem: Find domain of the function :

$$f(x) = \sqrt{x^2 - [x]^2}$$

Solution:

We analyze given function using its properties to find domain. Subsequently, we shall use graphical solution, which is more elegant. Now, for radical function,

$$\Rightarrow x^2 - [x]^2 \geq 0$$

Evaluation of this expression for integer values of x is easy. We know that [x] evaluates to x for all integer values of x :

$$[x] = x; \quad x \in Z$$

Squaring both sides,

$$[x]^2 = x^2; \quad x \in Z$$

$$x^2 - [x]^2 = 0; \quad x \in Z$$

However, evaluation of expression is slightly difficult for other values of x. Now, consider positive interval 1≤x<2. Here, [x] evaluates to 1 and its square is 1, which is less than or equal to x^2 . On the other hand, in negative interval -2≤x<-1, [x] evaluates to -2 and its square is 4, which is equal to or greater than x^2.

$$[x]^2 \leq x^2; \quad x > 0$$

$$[x]^2 \geq x^2; \quad x > 0$$

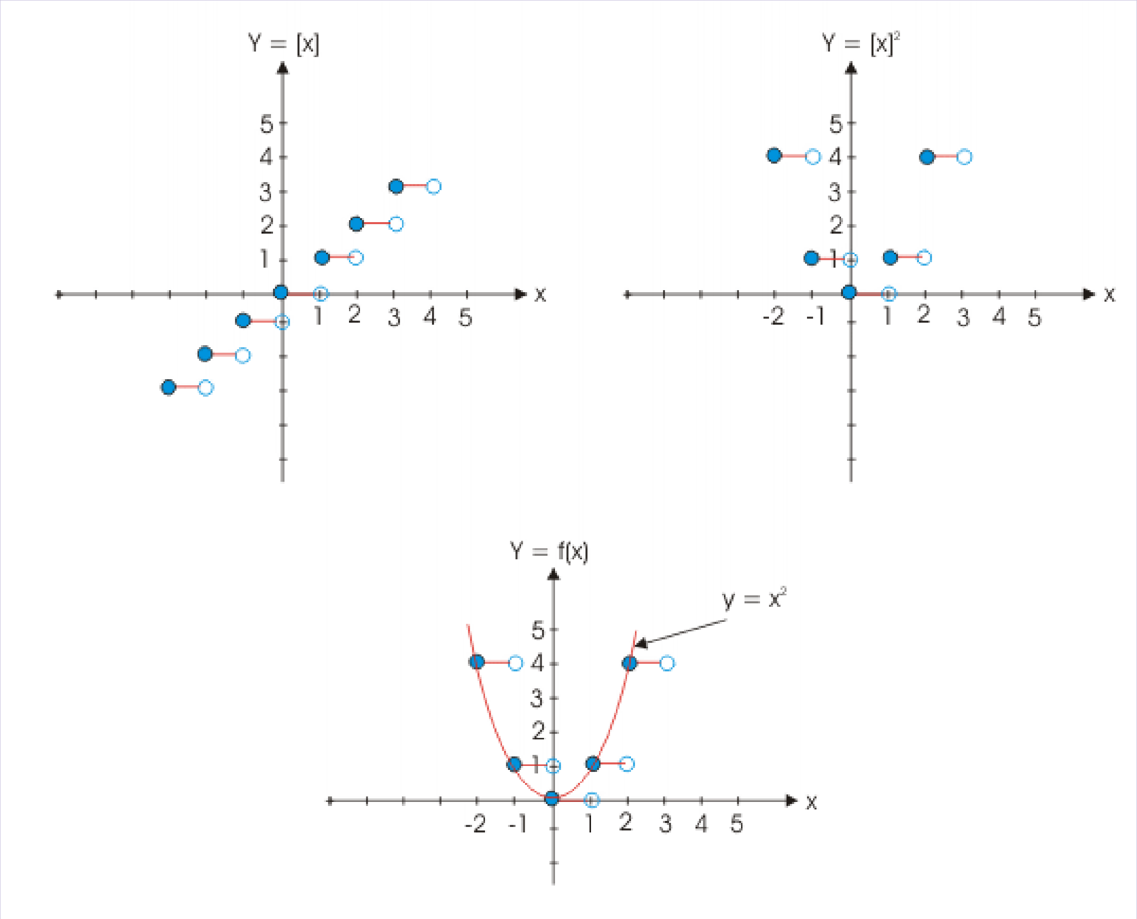
Note that we have included “equal to sign” for both intervals of x. Equal to sign is appropriate when x is integer. For x=0, expression evaluates to 0. It means expression is non-negative for all non-negative x. But expression also evaluates to 0 for negative integers. Hence, domain of given function is :

$$\text{Domain} = (0,\infty) \cup \{-n; n \in N\}$$

Graphical analysis

We draw $y = [x]$ and $y = [x]^2$ as in the first and second figures. Finally, we superimpose $y = x^2$ on the graph $y = [x]^2$ as shown in the third figure. Noting values of x for which value of x^2 is greater than or equal to $[x]^2$, the domain of the function is :

Domain



Domain is chosen for x such that difference of graphs is non-negative.

$$\text{Domain} = (0,\infty) \cup \{-n; n \in N\}$$

Exponential and logarithmic functions

Logarithmic and exponential functions are closely related functions. Logarithmic functions are useful in interpreting expressions/ equations, in which exponents are unknown. On the other hand, exponential functions are representation of natural process or mathematical relations, having exponential growth or decay. We shall encounter many expressions, involving these two functions in mathematics, while analyzing or describing processes. Further, two function types have simple derivatives and related integration results (we shall learn about derivatives and integration in calculus).

Symbolically, exponential and logarithmic functions, respectively, are :

$$f(x) = a^x$$

and

$$f(x) = \log_a x$$

In this module, we shall find out that exponential and logarithmic functions are inverse of each other. The roles of function "f(x)" and independent variable "x" are exchanged in two function definition. We shall clarify this point subsequently.

Exponential function

An exponential function relates every real number “x” to the exponentiation, “ a^x ”. In other words, we can say that an exponential function associates every real number (x) to a function given by :

$$f(x) = y = a^x$$

where :

1. The base “a” is positive real number, but excluding “1”. Symbolically, $a > 0$,$a \neq 1$.
2. The exponent “x” is a real number.

3. The number “y” represents the result of exponentiation, “ a^x ” and is a positive real number. Symbolically, y >0.

We need not memorize nature of "x","y" and "a". Rather we can investigate the same by reasoning. The base "a" can not be zero, because 0^x is not uniquely defined as required for invertible function. The exponential function is designed to be invertible. This means that "f(x)" and "x" be uniquely related. Also, if "a" is a negative number like "-2", then $(-2)^x$ may evaluate to positive or negative value depending on whether "x" is even or odd integer. Clearly sign of function depends on the nature of integer values of "x". In case, "x" is not an integer, then sign of $(-2)^x$ can not be interpreted for all values of "x". Further if a=1, then 1^x evaluates to "1" for all values of "x". Again function is not uniquely defined. In the nutshell, we conclude that base "a" is a positive number, but not equal to "1". In particular, the exponential function corresponding to base “e”, which is equal to 2.718281828, is called “natural” exponential function.

Once nature of "a" is decided, it is easy to find the nature of "y". Consider simplified exponents like 10^2 , 10^1 , 10^{-2}, 10^{-100}. All these numbers are greater than zero. This is true even if exponent is not integer. We conclude that "y" is positive number. Note that neither "a" nor "y" are zero. On the other hand, "x" by definition is a real number.

Clearly, the domain and range of exponential function are :

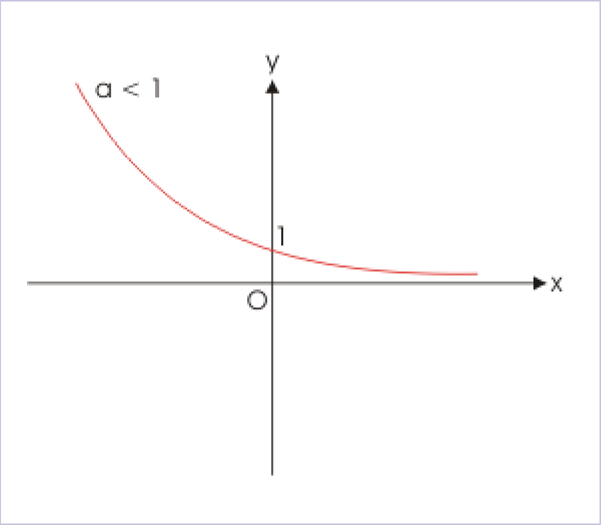
$$\text{Value of “x”} = \text{Domain} = R$$

$$\text{Value of “y”} = \text{Range} = (0,\infty)$$

We shall see subsequently that roles of “x” and “f(x)” are exchanged for logarithmic function. Here, “x” is the exponent i.e. the logarithmic value of a positive number and “f(x)” is the result of exponentiation, which is argument of logarithmic function. For this reason, we say that exponential and logarithmic functions are inverse to each other. As expected for inverse functions, we shall also see that domain and range of exponential and logarithmic functions are exchanged.

The nature of exponential function are different around a=1. The plots of exponential functions for two cases (i)0 < a < 1 and (ii) a > 1 are discussed here. If the base is greater than zero, but less than “1”, then the exponential function asymptotes to positive x-axis. It is easy to visualize the nature of plot. It is placed in the positive upper part as "f(x)" is positive. Also, note that $(0.25)^2$ is greater than ($(0.25)^4$. Hence, plot begins from a higher value to lower value as "x" increases, but never becomes equal to zero.

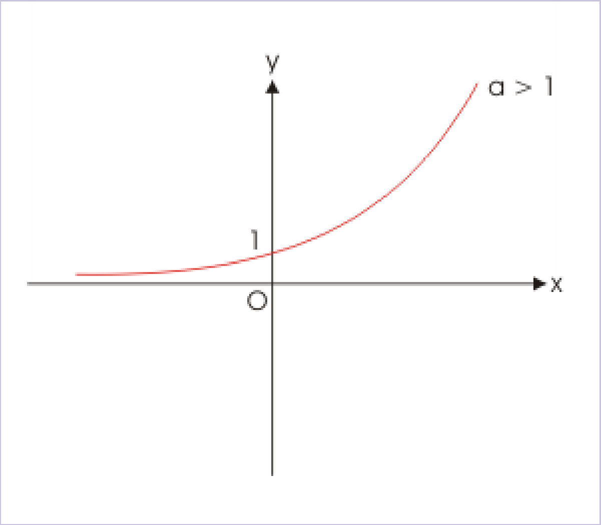
Exponential function



The plot of exponential function, when base is less than "1".

If the base is greater than “1”, then the exponential function asymptotes to negative x-axis. Again, it is easy to visualize the nature of plot. It is placed in the positive upper part as "f(x)" is positive. Also, note that $(1.25)^2$ is less than ($(1.25)^4$. Hence, plot begins from a lower value to higher value as "x" increases.

Exponential function



The plot of exponential function, when base is greater than "1".

Note that expanse of exponential function is along x – axis on either side of the y-axis, showing that its domain is R. On the other hand, the expanse of “y” is limited to positive side of y-axis, showing that its range is positive real number. Further, irrespective of base values, all plots intersect y-axis at the same point i.e. y = 1 as :

$$y = a^x = a^0 = 1$$

Logarithmic functions

A logarithmic function gives “exponent” of an expression in terms of a base, “a”, and a number, “x”. The following two representations, in this context, are equivalent :

$$a^y = x$$

and

$$f(x) = y = \log_a x$$

where :

1. The base “a” is positive real number, but excluding “1”. Symbolically, $a > 0$,$a \neq 1$.
2. The number “x” represents result of exponentiation, “ a^y ” and is also a positive real number. Symbolically, x >0.
3. The exponent “y” i.e. logarithm of “x” is a real number.

Note that neither “a” nor “x” equals to zero.

The expression of a logarithm for “x” on a certain base represents logarithmic function. In words, we can say that a logarithmic function associates every positive real number (x) to a real valued exponent (y), which is symbolically represented as :

$$f(x) = y = \log_a x; \quad a, x > 0, \quad a \neq 1$$

Following earlier discussion for the case of exponential function, we exclude "a = 1" as logarithmic function is not relevant to this base.

$$1^y = 1$$

We can easily see here that whatever be the exponent, the value of logarithmic function is “1". Hence, base “1” is irrelevant as exponent “y” is not uniquely associated with “x”.

From the defining values of "x" and "f(x)", we conclude that domain and range of logarithmic function is :

$$\text{Value of “x”} = \text{Domain} = (0,\infty)$$

$$\text{Value of “y”} = \text{Range} = R$$

Note that domain and range of logarithmic function is exchanged with respect to domain and range of exponential function.

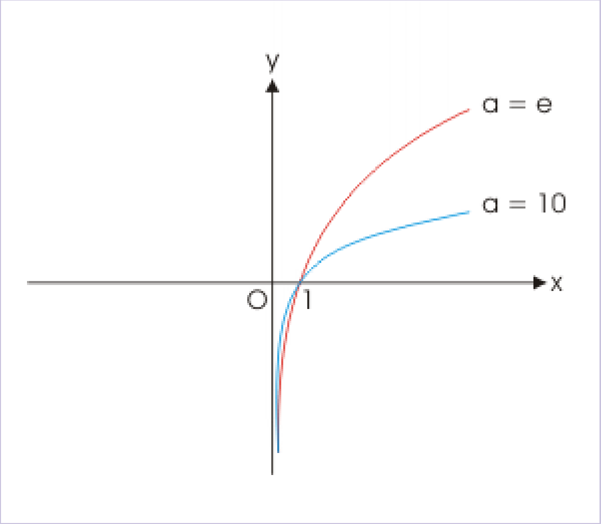
Base

The base of the logarithmic function can be any positive number. However, “10” and “e” are two common bases that we often use. Here, “e” is a mathematical constant given by :

$$e = 2.718281828$$

If we use “e” as the base, then the corresponding logarithmic function is called “natural” logarithmic function. The plots, here, show logarithmic functions for two bases (i) 10 and (ii) e.

Logarithmic function



The plots of logarithmic function on different bases.

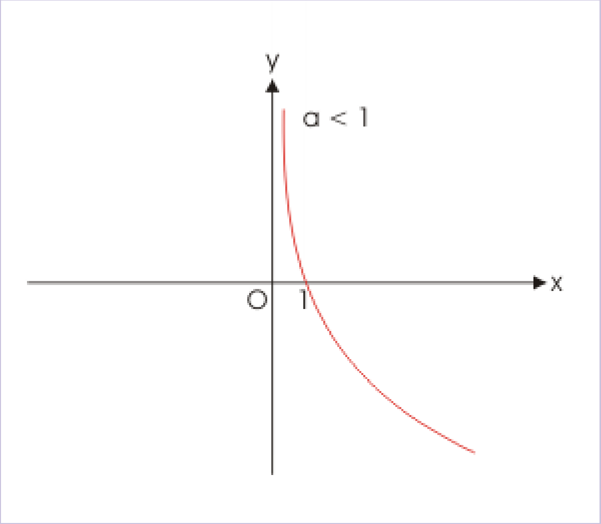
Note that expanse of logarithmic function "f(x)" is along y – axis on either side of axis, showing that its range is R. On the other hand, the expanse of “x” is limited to positive side of x-axis, showing that domain is positive real number. Further, irrespective of bases, logarithmic plots intersect x-axis at the same point i.e. x = 1 as :

$$x = a^y = a^0 = 1$$

Graphs

We have noted the importance of base “1” for logarithmic function. Base “1” plays an important role for determining nature of logarithmic function. Here, we have drawn plots for two cases : (i) 0<a < 1 and (ii) a > 1. If the base is greater than zero, but less than “1”, then the logarithm function asymptotes to positive y-axis. Since, "x" is positive number, plot falls only in the right side quadrants. Also, note that $\log_{0.1} 0.001 = 3$ is less than $\log_{0.1} 0.01 = 2$. Hence, plot begins from a higher value to lower value as "x" increases. Further, $\log_{0.1} 10 = -1$. We deduce that plot moves below x-axis for "x" greater than "1".

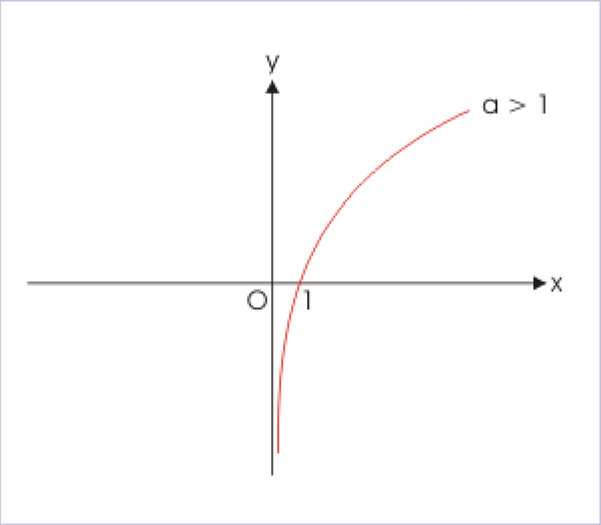
Logarithmic function



The plot of logarithmic function, when base is less than "1".

If the base is greater than “1”, then the logarithm function asymptotes to negative y-axis. Since, "x" is positive number, plot again falls only in the right side quadrants. Also, note that $\log_{10} 10 = 1$ is greater than $\log_{10} 100 = 2$. Hence, plot begins from a lower value to greater value as "x" increases. Further, $\log_{10} 10 = 1$. We deduce that plot moves above x-axis for "x" greater than "1".

Logarithmic function



The plot of logarithmic function, when base is greater than "1".

Logarithmic identities

Some of the important logarithmic identities are given here without proof. Idea, here, is to simply equip ourselves so that we can work with logarithmic functions in conjunction with itself or with other functions.

$$\log_a a = 1$$

$$\log_a 1 = 0$$

$$x = a^y = a^{\log_a x}$$

$$a^{\log_b c} = c^{\log_b a}$$

$$\log_a(xy) = \log_a(x) + \log_a(y)$$

$$\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$$

$$\log_a(x^y) = y\log_a(x)$$

$$\log_a\left(x^{\frac{1}{y}}\right) = \frac{\log_a(x)}{y}$$

Change of base

Sometimes, we are required to work with logarithmic expressions of different bases. In such cases, we convert them to same base, using following relation :

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$$

Example:
Problem : Find the domain of the function given by :

$$f(x) = \frac{1}{\log_{10}(1-x)}$$

Solution : The given function is reciprocal of a logarithmic function. Therefore, we first need to ensure that logarithmic function does not evaluate to zero for a value of "x". A logarithmic function evaluates to zero if its argument is equal to "1". Hence,

$$1 - x \neq 1$$

$$\Rightarrow x \neq 0$$

Further domain of logarithmic function is a positive real number. For that :

$$1 - x > 0$$

$$\Rightarrow 1 > x$$

$$\Rightarrow x < 1$$

Combining two results, the domain of the given function is :

$$\text{Domain of “f”} = (-\infty, 1) - \{0\}$$

Example:

Problem : Find all real values of “x” such that :

$$1 - e^{\frac{1}{x} - 1} > 0$$

Solution : Solving for the exponential function, we get following inequality :

$$\Rightarrow e^{\frac{1}{x} - 1} < 1$$

Now, we know that domain of an exponential function is “R”. However, this information is not helpful here to find values of “x” that satisfies the inequality. Taking natural logarithm on either side of the equation,

$$\Rightarrow \frac{1}{x} - 1 < \log_e 1$$

But, logarithm of “1” i.e. $\log_e 1$ is zero. Hence,

$$\Rightarrow \frac{1}{x} - 1 < 0$$

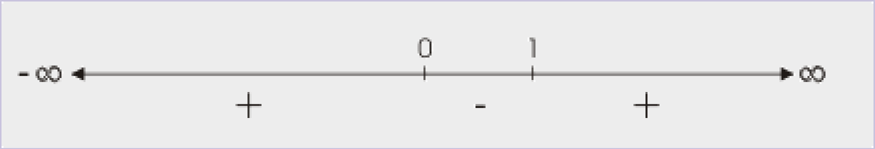
$$\Rightarrow \frac{1 - x}{x} < 0$$

This is rational inequality. In order to solve this inequality, we need to multiply each side of the inequality by -1 so that "x" in "1-x" becomes positive. This multiplication changes the inequality to "greater than" inequality.

$$\Rightarrow \frac{x - 1}{x} > 0$$

Here, critical points are 0 and 1. The rational expression is positive in the intervals on either side of middle interval.

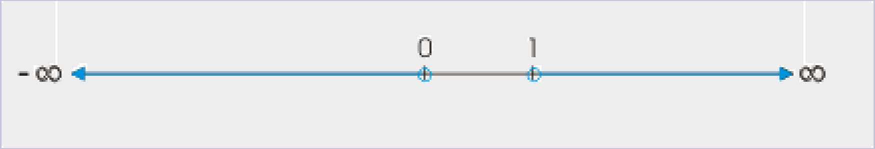
Sign convention



Positive and negative intervals.

Picking the intervals for which expression is positive :

Intervals



Intervals for which rational function is positive.

$$\text{Domain} = (-\infty, 0) \cup (1,\infty)$$

Logarithmic inequality

The nature of logarithmic function is dependent on the base value. We know that base of a logarithmic function is a positive number excluding “1”. The value of “1” plays important role in deciding nature of logarithmic function and hence that of inequality associated to it. Let us consider an equality :

$$\log_a x > y$$

What should we conclude : $x > a^y$ or $x < a^y$? It depends on the value of “a”. We can understand the same by considering LHS of the inequality equal to an exponent z :

$$\log_a x = z$$

If a > 1, then "a^z" will yield a greater “x” than "a^y", because z>y (it is given by the inequality). On the other hand, if 0<a<1, then "a^z" will yield a smaller “x” than "a^y", because z>y. We can understand this conclusion with the help of an example. Let

$$\log_2 x > 3$$

Let

$$\log_2 x = 4$$

Clearly, $2^4 > 2^3 \quad \text{as} \quad 16 > 8$

Let us now consider a <1,

$$\log{}_{0.5} x > 3$$

Let

$$\log_{0.5} x = 4$$

Clearly, $0.5^4 < 0.5^3 \quad \text{as} \quad 0.0625 < 0.125$. Thus, we finally conclude :

$$\log_a x > y \Leftrightarrow x > a^y; \quad a > 1$$

$$\log{}_a x > y \Leftrightarrow x < a^y; \quad 0 < a < 1$$

We have used two ways notation to indicate that interpretation of logarithmic inequality is true in either direction. Similarly, we can conclude that :

$$\log_a x < y \Leftrightarrow x < a^y; \quad a > 1$$

$$\log{}_a x < y \Leftrightarrow x > a^y; \quad 0 < a < 1$$

Logarithmic inequality involving logarithmic functions on same base on either side of the inequality can be interpreted as :

$$\log_a x > \log_a y \Leftrightarrow x > y; \quad a > 1$$

$$\log_a x > \log_a y \Leftrightarrow x < y; \quad 0 < a < 1$$

Alternate method

We know that if x>y, then :

$$a^x > a^y; \quad a > 1 \quad \text{and} \quad a^x < a^y; \quad 0 < a < 1$$

Let us now consider the inequality $\log_a x > y$. Following above deduction, we use the term of inequality as powers on a common base :

$$a^{\log_a x} > a^y; \quad a > 1 \quad \text{and} \quad a^{\log_a x} < a^y; \quad 0 < a < 1$$

Using identity, $x = a^{\log_a x}$,

$$x > a^y; \quad a > 1 \quad \text{and} \quad x < a^y; \quad 0 < a < 1$$

Similarly, we can deduce other results of logarithmic inequalities.

Domain and range of exponential and logarithmic function

Working rules : We shall be using following definitions/results for solving problems in this module :

- $y = \log_a x, \quad \text{where} \quad a > 0, a \ne 1, x > 0, y \in R$
- $y = \log_a x \Leftrightarrow x = a^y$
- $\text{If} \quad \log_a x \ge y, \quad \text{then} \quad x \ge a^y, \quad \text{if} \quad a > 1$
- $\text{If} \quad \log_a x \ge y, \quad \text{then} \quad x \le a^y, \quad \text{if} \quad a < 1$

Domain of different logarithmic functions

Example:

Problem : Find the domain of the function given by (Be aware that "x" appears as base of given logrithmic function):

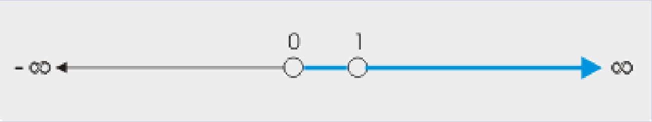
$$f(x) = \log_x 2$$

Solution : By definition of logarithmic function, we know that base of logarithmic function is a positive number excluding x =1.

$$x > 0, \quad x \ne 1$$

Hence, domain of the given function is :

Domain of the function



Thick line represents domain of the given function.

$$\text{Domain} = (0,\infty) - \{1\}$$

or,

$$\text{Domain} = (0,1) \cup \{1,\infty\}$$

Example:

Problem : Find the domain of the function given by :

$$f(x) = \log_{10} \frac{x^2 - 5x + 6}{x^2 + 5x + 9}$$

Solution : The argument (input to the function) of logarithmic function is a rational function. We need to find values of “x” such that the argument of the function evaluates to a positive number. Hence,

$$\Rightarrow \frac{x^2 - 5x + 6}{x^2 + 5x + 9} > 0$$

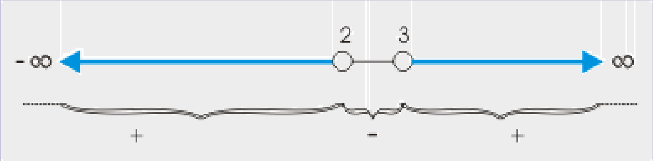
In this case, we can not apply sign scheme for the rational function as a whole. Reason is that the quadratic equation in the denominator has no real roots and as such can not be factorized in linear factors. We see that discreminant,"D", of the quadratic equation in the denominator, is negative :

$$\Rightarrow D = b^2 - 4ac = 5^2 - 4X1X9 = 25 - 36 = -11$$

The quadratic expression in denominator is positive for all value of x as coefficient of squared term is positive. Clearly, sign of rational function is same as that of quadratic expression in the numerator. The coefficient of squared term of the numerator “x^2”, is positive for all values of “x”. The quadratic expression in the numerator evaluates to positive for intervals beyond root values. The roots of the corresponding equal equation is :

$$\Rightarrow x^2 - 2x - 3x + 6 = 0 \quad \Rightarrow x(x-2) - 3(x-2) = 0 \quad \Rightarrow (x-2)(x-3) = 0$$

Domain of the function



Thick line represents domain of the given function.

$$x < 2 \quad \text{or} \quad x > 3$$

$$\text{Domain} = (-\infty, 2) \quad \cup \quad (3,\infty)$$

Example:

Problem : Find the domain of the function given by :

$$f(x)=\sqrt{\log_{10}\frac{6x-x^2}{8}}$$

Solution : The function is a square root of a logarithmic function. On the other hand argument of logarithmic function is a rational function. In order to find the domain of the given function, we first determine what values of “x” are valid for logarithmic function. Then, we apply the condition that expression within square root should be non-negative number. Domain of given function is intersection of intervals of x obtained for each of these conditions. Now, we know that argument (input to function) of logarithmic function is a positive number. This implies that we need to find the interval of “x” for which,

$$\Rightarrow \frac{6x-x^2}{8}>0$$

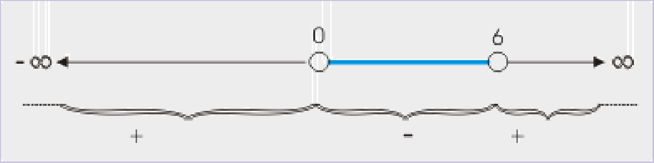
$$\Rightarrow 6x-x^2>0$$

In above step, we should emphasize here that we multiply “8” and “0” and retain the inequality sign because 8>0. Now, we multiply the inequality by “-1”. Therefore, inequality sign is reversed.

$$\Rightarrow x^2-6x<0$$

Here, roots of corresponding quadratic equation “ x^2-6x” is x = 0, 6. It means that middle interval between “0 and 6” is negative as coefficient of “x^2” is positive i.e. 6>0. Hence, interval satisfying the inequality is :

Domain of the function



Thick line represents domain of the given function.

$$0<x<6$$

Now, we interpret second condition according to which the whole logarithmic expression within the square root should be a non-negative number.

$$\Rightarrow \log_{10}\frac{6x-x^2}{8} \geq 0$$

We use the fact that if $\log_a x \geq y$, then $x \geq a^y$ for $a > 1$. This gives us the inequality as given here,

$$\Rightarrow \frac{6x-x^2}{8} \geq 10^0$$

$$\Rightarrow \frac{6x-x^2}{8} \geq 1$$

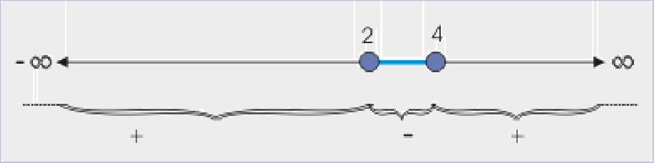
$$\Rightarrow 6x-x^2 \geq 8 \quad \Rightarrow 6x-x^2-8 \geq 0$$

$$\Rightarrow x^2-6x+8 \leq 0 \quad \Rightarrow x^2-2x-4x+8 \leq 0$$

$$\Rightarrow x(x-2)-4(x-2) \leq 0 \quad \Rightarrow (x-2)(x-4) \leq 0$$

Clearly, “2” and “4” are the roots of the corresponding quadratic equation. Following sign scheme, we pick middle negative interval :

Domain of the function

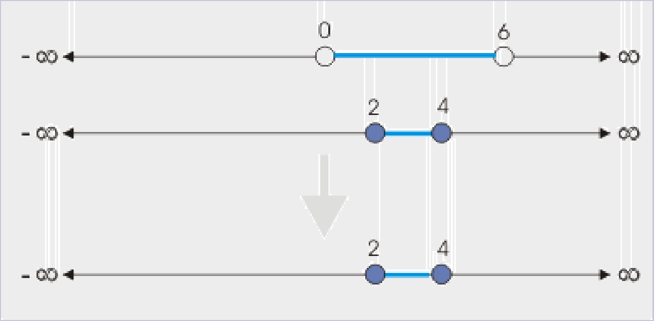


Thick line represents domain of the given function.

$$2 \leq x \leq 4$$

Now, the interval of “x” valid for real values of “f(x)” is the one which satisfies both conditions simultaneously i.e. the interval common to two intervals determined. Hence,

Domain of the function



Thick line represents domain of the given function.

$$\text{Domain} = 0 < x < 6 \quad \cap \quad 2 \leq x \leq 4$$

$$\text{Domain} = 2 \leq x \leq 4 = [2,4]$$

Example:

Problem : Find the domain of the function given by :

$$f(x) = \sqrt{\left\{(\log_{0.2} x)^3 + \log_{0.2} x^3 X \log_{0.2} 0.0016x + 36\right\}}$$

Solution : The function is square root of an expression, consisting logarithmic functions. Here, we first need to simplify expression, using logarithmic identities, before attempting to find domain of the function. Let us first simplify the middle term of the given expression, using logarithmic identities :

$$\log_{0.2} x^3 X \log_{0.2} 0.0016x = 3 \log_{0.2} x X \log_{0.2} 0.2^4 x$$

$$\Rightarrow \log_{0.2} x^3 X \log_{0.2} 0.0016x = 3 \log_{0.2} x X (4 \log_{0.2} 0.2 + \log_{0.2} x)$$

We observe that all logarithmic functions have the base of “0.2”. Let us consider that $z = \log_{0.2} x$, then logarithmic expression within square root is :

$$\Rightarrow z^3 + 3z(4+z) + 36 = z^3 + 3z^2 + 12z + 36 = z^2(z+3) + 12(z+3)$$

$$\Rightarrow z^3 + 3z(4+z) + 36 = \left(z^2 + 12\right)(z+3)$$

Now, this expression is non-negative for square root to be real. Hence,

$$\Rightarrow \left(z^2 + 12\right)(z+3) \geq 0$$

But, we see that $z^2 + 12$ is a positive number as term z^2 is positive. It means that :

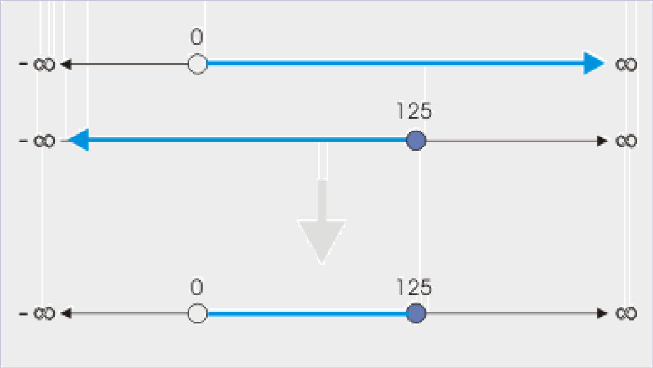
$$\Rightarrow (z+3) \geq 0 \quad \Rightarrow \log \log_{0.2} x \geq -3 \quad \Rightarrow x \leq 0.2^{-3} \quad \Rightarrow x \leq \frac{1}{0.008} \quad \Rightarrow x \leq 125$$

Note that we have reversed the inequality as the base is 0.2, which is less than 1. Further, we have substituted as :

$$z = \log_{0.2} x$$

This logarithmic function is valid by definition for all positive value of “x”. Now, the domain of given function is the intersection of two intervals as shown in the figure.

Domain of the function



Thick line represents domain of the given function.

$$\text{Domain} = (0,125]$$

Range of logarithmic function

Example:

Problem : Find range of the function :

$$f(x) = \frac{e^x - e^{-|x|}}{e^x + e\,|x|}$$

Solution : We observe that for x≤0,

$$f(x) = 0$$

For x>0

$$\Rightarrow y = f(x) = \frac{e^x - e^{-|x|}}{e^x + e\,|x|} = \frac{e^{2x} - 1}{2e^{2x}}$$

$$\Rightarrow yX2e^{2x} = e^{2x} - 1 \quad \Rightarrow (1-2y)e^{2x} = 1$$

$$\Rightarrow e^{2x} = \frac{1}{1-2y}$$

We can see that $e^{2x} \geq 1$ for all x. Hence,

$$\Rightarrow \frac{1}{1-2y} \geq 1 \quad \Rightarrow \frac{1}{1-2y} - 1 \geq 0 \quad \Rightarrow \frac{1-1+2y}{1-2y} \geq 0$$

$$\Rightarrow \frac{2y}{1-2y} \geq 0$$

$$\Rightarrow \frac{2y}{2y-1} \leq 0$$

Here, critical points are 0,1. Thus, range of the given function is :

$$\text{Range} = \left[0, \frac{1}{2}\right)$$

Exercise

Exercise:

Problem: Find the domain of the function given by :

$$f(x) = 2^{\sin^{-1}(x)}$$

Solution:

The exponent of the exponential function is inverse trigonometric function. Exponential function is real for all real values of exponent. We see here that given function is real for the values of “x” corresponding to which arcsine function is real. Now, domain of arcsine function is [-1,1]. This is the interval of "x" for which arcsine is real. Hence, domain of the given function, “f(x)” is :

$$\text{Domain} = [-1,1]$$

Exercise:

Problem: Find the domain of the function given by :

$$f(x) = \log_{10}\left\{\sqrt{(8-x)} + \sqrt{(x-2)}\right\}$$

Solution:

The argument (input to the function) of logarithmic function is addition of two square roots. We need to find values of "x" such that the argument of the logarithmic function evaluates to a positive number. An unsigned square root is a positive number by definition. It can not be negative. Symbolically, √x is a positive number. Clearly, each of the square roots is a positive number. Hence, their addition is also a positive number. Thus, we see that the requirement of the argument of a logarithmic function being a positive number, is automatically fulfilled by virtue of the property of an unsigned square root.

We, therefore, only need to evaluate "x" for which each of the square roots is real. In other words, the expressions in each of the square roots is a non-negative integer.

$$8 - x \geq 0 \quad \Rightarrow x - 8 \leq 0 \quad \Rightarrow x \leq 8$$

$$x - 2 \geq 0 \quad \Rightarrow x \geq 2$$

The two square root functions are added to form the argument of logarithmic function. We know that domain of function resulting from addition is intersection of domains of individual square root function. Hence,

$$\text{Domain} = [2,8]$$

Exercise:

Problem: Find the domain of the function :

$$f(x) = \log_{10}\left\{1 - \log\left(x^2 - 3x + 12\right)\right\}$$

Solution:

Hints : There are two logarithmic functions composing the given function. Let us call them outer and inner. For outer logarithmic function,

$$\Rightarrow 1 - \log\left(x^2 - 3x + 12\right) > 0$$

$$\Rightarrow \log\left(x^2 - 3x + 12\right) < 1$$

$$\Rightarrow \log_{10}\left(x^2 - 3x + 12\right) < \log_{10} 10$$

$$\Rightarrow x^2 - 3x + 12 < 10$$

$$\Rightarrow x^2 - 3x + 2 < 0$$

$$\Rightarrow (x - 1)(x - 2) < 0$$

$$\Rightarrow x \in (1,2)$$

For inner logarithmic function,

$$x^2 - 3x + 12 > 0$$

Here, coefficient of squared term is positive and and D<0. Hence, this inequality is true for all real x i.e. x R. Now, domain of given function is intersection of two intervals.

$$\text{Domain} = (1,2)$$

Exercise:

Problem: Problem 3 : Find the domain of the function given by :

$$f(x) = \log_2 \log_3 \log_4 x$$

Solution:

The function is formed by nesting three logarithmic functions. Further base of logarithmic functions are different. For determining domain we (i) find value of "x" for which "$\log_4 x$" is real (ii) find range of "$\log_4 x$" for which "$\log{}_3(\log{}_4 x)$" is real and (iii) find range of "$\log{}_3(\log{}_4 x)$" for which f(x) is real.

For "$\log_4 x$" to be real, x is a positive number. It means,

$$x > 0$$

For "$\log{}_3(\log{}_4 x)$" to be real, "$\log_4 x$" is required to be positive. It means,

$$\log{}_4 x > 0$$

Using the fact that if $\log{}_a x \geq y, \quad \text{then} \quad x \geq a^y \quad \text{for} \quad a > 1$, we have :

$$\Rightarrow x > 4^0 \quad \Rightarrow x > 1$$

For “f(x)” to be real, “$\log_3(\log_4 x)$” is required to be positive. It means,

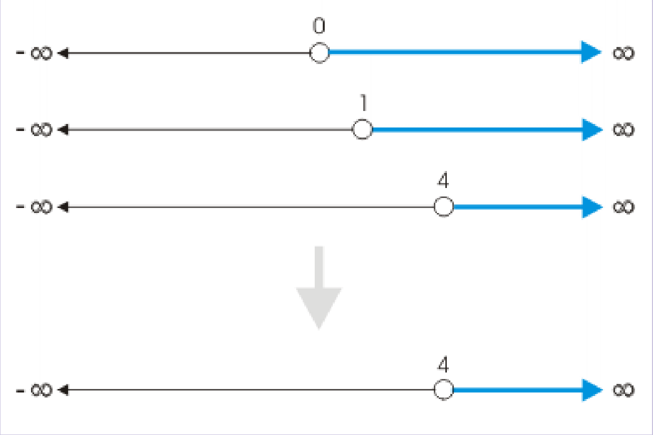
$$\Rightarrow \log_3(\log_4 x) > 0$$

$$\Rightarrow \log_4 x > 1$$

$$\Rightarrow x > 4^1 \quad \Rightarrow x > 4$$

Combining three intervals so obtained,

Domain of the function



Thick line represents domain of the given function.

$$\text{Domain} = (4,\infty)$$

Exercise:

Problem: Find the range of the function :

$$f(x) = \log_{10}\left(x^2 - 3x + 4\right)$$

Solution:

Hints : We need to find minimum and maximum value of logarithmic function for the values of x in domain of the function. The argument of logarithmic function is a quadratic function, whose coefficient of squared term is positive and D <0. It means its graph is a parabola opening up in the positive side of y-axis. The minimum value of the quadratic expression is :

$$y_{\min} = -\frac{D}{4a}$$

$$y_{\max} = \infty$$

Now, we know that graph of logarithmic function for base, a > 1, is a continuously increasing graph. It means that value of logarithmic function, corresponding to min and max values of quadratic expression is the range of given function.

$$\Rightarrow f\left(\frac{7}{4}\right) = \log 10\left(\frac{7}{14}\right)$$

$$\Rightarrow f(x \to \infty) \to \infty$$

Hence, range of given function is :

$$\text{Range} = \left(\log 10\left(\frac{7}{14}\right), \infty\right)$$

Exercise:

Problem: Find domain and range if

$$e^x - e^{f(x)} = e$$

Solution:

Rearranging, we have :

$$\Rightarrow e^{f(x)} = e^x - e$$

Taking logarithm on either sides of equation,

$$\Rightarrow y = f(x) = \log(e^x - e)$$

For logarithmic function,

$$\Rightarrow e^x - e > 0 \quad \Rightarrow e^x > e \quad \Rightarrow x > 1$$

$$\text{Domain} = (1,\infty)$$

In order to find range, we solve function expression for y. In exponential form,

$$\Rightarrow e^y = e^x - e \quad \Rightarrow e^x = e^y - e$$

Taking logarithm on either sides of equation,

$$\Rightarrow x = \log_e(e^y - e)$$

For logarithmic function,

$$\Rightarrow e^y - e > 0 \quad \Rightarrow e^y > e \quad \Rightarrow y > 1$$

$$\text{Range} = (1,\infty)$$

Trigonometric functions

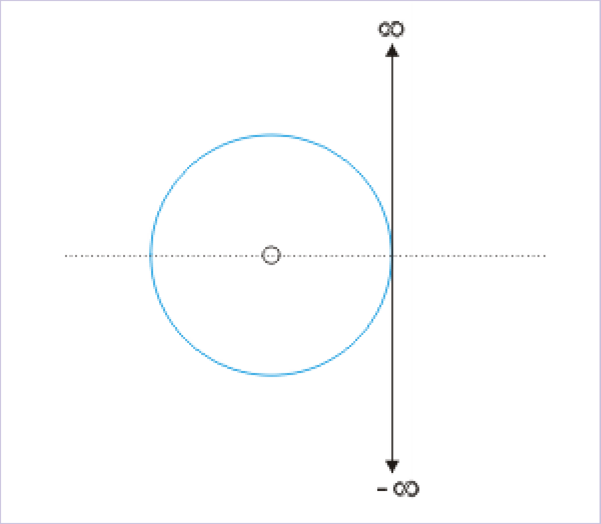
We are familiar with trigonometric ratios, identities and their applications. In this module, we shall revisit the concept of trigonometric ratios from the perspective of a function. For this, we shall first recapitulate a bit of basics and important results and then emphasize: how can we conceive trigonometric ratio as a function?

The most important aspect, here, is the extension of the concept of angle beyond the domain of 2π i.e. the angle of "1" revolution. This concept is followed by the investigation of trigonometric ratios, which is originally defined for acute angle. Here, we shall apply these ratios in the context of any real value angle, represented on a real number line.

Angle and real number

The measurement of angle is constrained to a circular periphery. We can unwind this constraint and think of angle as a real number, extending from minus infinity to plus infinity. For this, we imagine the circular periphery straightened into a line. Alternatively, we may think angle be represented along a straight line like real number and then think to bend straight line along the periphery of the circle. Following this visualization, we consider angle as if it were represented by a real number line, which is tangent to the circle.

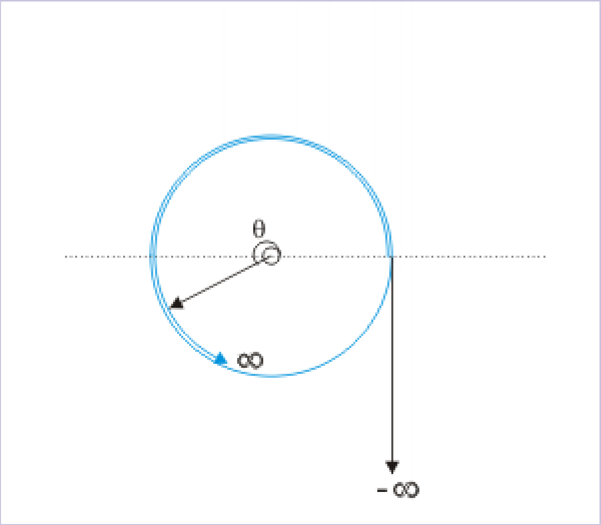
Angle



Angle and real number.

The positive section of the real number line can be wrapped many times over in the anticlockwise direction. Similarly, the negative section of the number line can be wrapped many times over in the clockwise direction.

Angle



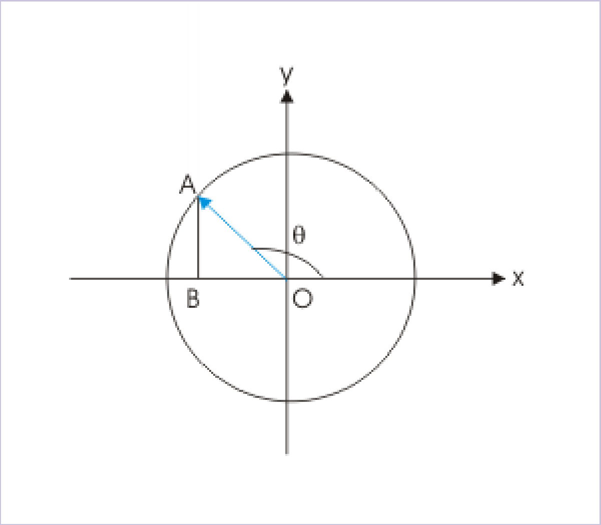
Angle and real number.

We consider representation of angle on real number line equivalent to measurement of angle from a reference direction about the central vertex “O” in as many revolutions as required. The measurement of angle in anticlockwise direction is considered positive and negative in clockwise direction.

Trigonometric ratios

Trigonometric ratios are defined for acute angle in a right angle triangle. Even for angles, which are not acute, we consider trigonometric ratios as ratios of sides or ratios of a side and hypotenuse of the right angle triangle OAB, which is constructed with the terminal ray, “OA” (measuring angle from the initial position in x-direction) and x-axis. The cosine of angle “θ”, in terms of side and hypotenuse of triangle OAB, is :

Trigonometric ratio



Trigonometric ratio of an angle greater than acute angle.

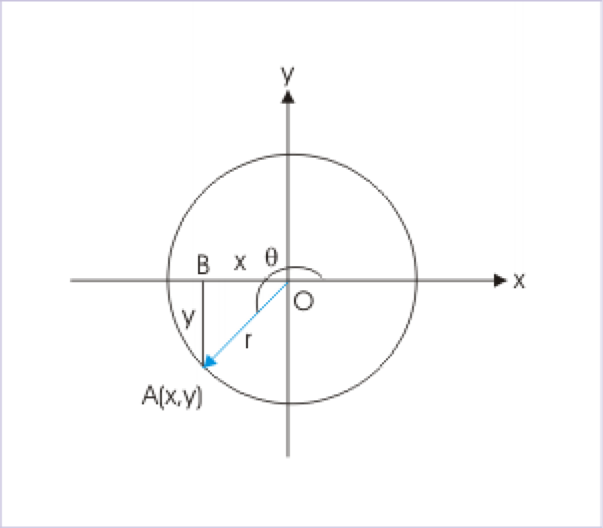
$$\cos\theta = \frac{OB}{OA}$$

Clearly, the sign of ratio is given by the sign of the side of the right angle triangle OAB, involved in the ratio. We attach sign to bidirectional measurements along x and y axes. We can not attach sign to the radial ray OA as it can be directed in multiple directions. In the case shown above, side of the triangle “OB” is negative with respect to positive x-direction. As such, the cosine of “θ” in this particular case is negative. However, note that “AB” is positive and hence sine of the angle, which involves “AB”, is positive for the same angle.

$$\sin\theta = \frac{AB}{OA}$$

Alternatively, the sign of “x” and “y” coordinates of the final ray on the circle decides the sign of trigonometric ratio. As one of the coordinates is involved in the ratio, its sign becomes the sign of trigonometric ratio. Consider the position, “A”, shown in the figure.

Sing of trigonometric ratio



The sign of trigonometric ratio is decided by the sign of coordiantes.

$$\sin\theta = \frac{y}{r} < 0 \quad \text{as “y” is negative}$$

$$\cos\theta = \frac{x}{r} < 0 \quad \text{as “x” is negative}$$

$$\tan\theta = \frac{y}{x} > 0 \quad \text{as both “x” and “y” are negative}$$

Unit circle

The angle and ratios defined in reference with circle is independent of the size of circle i.e. its radius. If radius is considered to be “1”, then we link trigonometric ratios directly to the coordinates of the tip of the terminal ray. Let x,y be the coordinates of a point “A” on a unit circle. Then,

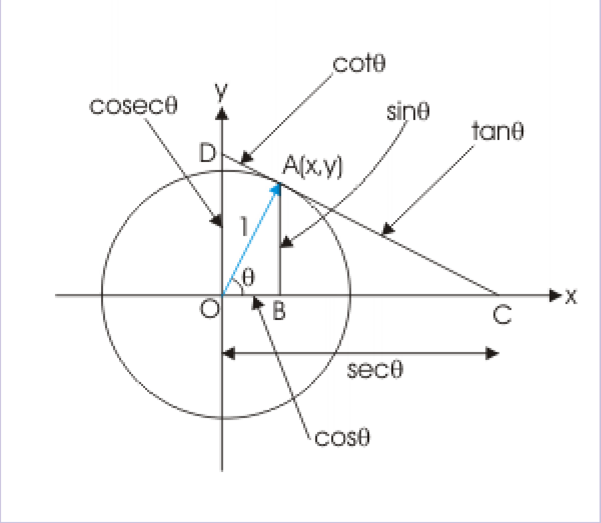
$$\Rightarrow \sin\theta = \frac{y}{r} = \frac{y}{1} = y$$

$$\Rightarrow \cos\theta = \frac{x}{r} = \frac{x}{1} = x$$

$$\Rightarrow \tan\theta = \frac{y}{x}$$

The figure below shows what these trigonometric ratios mean with reference to circle, tangent and coordinates.

Trigonometric ratios



Geometric meaning of trigonometric ratios on unit circle.

Trigonometric functions

There are six of trigonometric ratios. In the following sub-sections, we describe each trigonometric function with corresponding domain, range and graph. In particular, we shall come to know that some of these trigonometric functions are not defined for all values of angles. Further, we shall deliberately denote angle by variable “x” – not by “θ” as conventionally denoted. This is to emphasize that angle is a real number.

Besides, domain and range, we shall also discuss periodicity and polarity of each trigonometric function. We refer a function periodic if its values are repeated after certain interval. Graphically, periodic function has a fundamental segment, which can be used to draw plot of the function by repeating that fundamental segment again and again. Mathematically, we say that f(x+T) = f(x), where T is fundamental period.

Here, we shall make use of one important rule about periodic function. If T is the period of function f(x), then period of function $af(kx \pm b)$ is $\frac{T}{|k|}$, where a,b and k are real numbers. Important points to note that a and b do not affect period, but coefficient of x i.e. k affect period and is given by $\frac{T}{|k|}$.

On the other hand, polarity refers to whether the function is even or odd. If f(x) = f(-x), then function is even and its plot is symmetric about y-axis. If f(x) = -f(x), then function is odd and its plot is symmetric about origin.

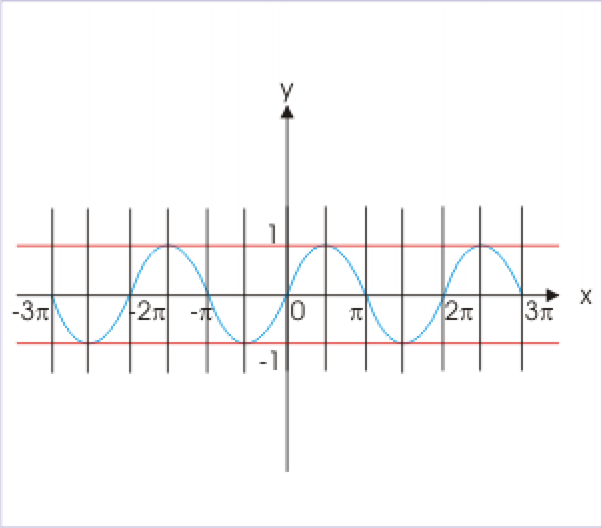
Sine function

For each real number “x”, there is a sine function defined as :

$$f(x) = \sin(x)$$

The plot of sin(x) .vs. x is shown here.

Sine function



Graph of sine function

The plot, here, is continuous and period is "2π". Think period of the function in term of minimum segment which can be used to extend the plot on either side. Further as sin(-x) = -sinx, sine function is an odd function. This fact is also substantiated by the fact that plot is symmetric about origin - not y-axis.

Since function holds for all values of “x”, its domain is “R”. On the other hand, the values of sine function is bounded between “-1” and “1”, inclusive of end points. Hence, domain and range of sine function are :

$$\text{Domain} = R$$

$$\text{Range} = [-1,1]$$

Let us now consider sine function which is given as :

$$f(x) = A\sin(x)$$

Multiplying sine function by a constant A does not change the periodicity of function. However, it changes the maximum and minimum values of the function. The plot extends from -A to A along y-axis as against from -1 to 1 when function is not multiplied by a constant. This, in turn, changes the range of the function :

$$\text{Range} = [-\text{A},\text{A}]$$

We now consider yet another form of sine function which is given as :

$$f(x) = A \sin(kx)$$

Multiplying argument x of sine function by a constant k does not change the nature of plot. However, it changes the periodicity of the function. Recall that if T is the period of function f(x), then period of function $af(kx \pm b)$ is $\frac{T}{|k|}$ Clearly, the period of sin(kx) is $\frac{T}{|k|}$. If |k| is less than 1, then period is more than 2π and if |k| is greater than 1, then period is less than 2π.

Example:

Problem : Find domain and range of function :

$$f(x) = \sin x + 2$$

Solution : We know that domain of sinx is real number set R and range is [-1,1]. The given function is real for all real values of x. Hence, its domain remains R. On the other hand, minimum and maximum values of function changes from that corresponding to sinx function :

$$y_{\text{min}} = -1 + 2 = 1$$

$$Y_{\text{max}} = 1 + 2 = 3$$

Hence, range of given function is [1,3]. It is evident that graph of function is that of graph of sinx shifted up by 2 units.

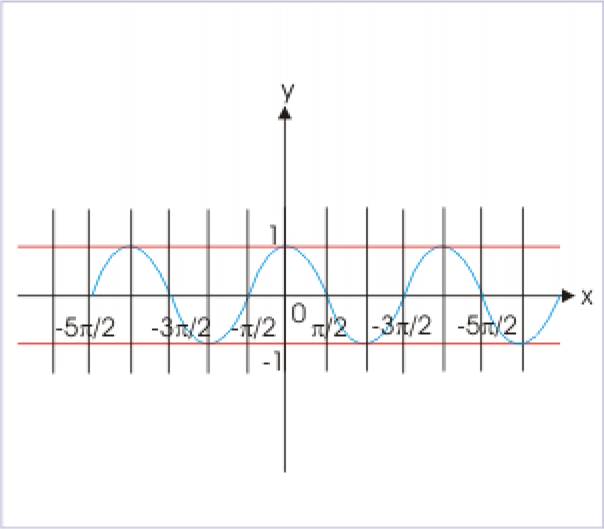
Cosine function

For each real number “x”, there is a cosine function defined as :

$$f(x) = \cos(x)$$

The plot of cos(x) .vs. x is shown here.

Cosine function



Graph of Cosine function

The plot, here, is continuous and period is "2π". Think period of the function in term of minimum segment which can be used to extend the plot on either side. Further as cos(-x) = cosx, cosine function is an even function. This fact is also substantiated by the fact that plot is symmetric about y-axis.

Since function holds for all values of “x”, its domain is “R”. On the other hand, the values of cosine function is bounded between “-1” and “1”, inclusive of end points. Hence, domain and range of sine function are :

$$\text{Domain} = R$$

$$\text{Range} = [-1,1]$$

When cosine function is given as f(x) = Acosx, maximum and minimum values of function becomes -A and A. The range is modified as :

$$\text{Range} = [-\text{A},\text{A}]$$

When cosine function is given as f(x) = Acos(kx), the period of cosine function is given by $\frac{T}{|k|}$.

Example:
Problem : Find domain range of the function :

$$f(x) = 12\sin x + 5\cos x$$

Solution : The given function comprises of sine and cosine functions. Here, we reduce given function in terms of one trigonometric function and then find range of the function. This reduction is required as otherwise it would be difficult to estimate when two trigonometric functions together evaluates to minimum and maximum values. Let us put,

$$a\cos\alpha = 12$$

$$a\sin\alpha = 5$$

Clearly, $a = \sqrt{12^2 + 5^2} = 13$. Putting these values/ expression in function,

$$f(x) = 13(\cos\alpha\sin x + \sin\alpha\cos x) = 13\sin(x+\alpha)$$

We know that range of sine function is [-1,1]. Hence, range of given function is :

$$\text{Range}[-13, 13]$$

Tangent function

For a real number “x”, there is a tangent function defined as :

$$f(x) = \tan(x)$$

Note that defining statement defines the function for a real number “x” – not for "each" real “x” as in the case of sine and cosine functions. It is so because, tangent function is not defined for all real values of “x”. Let us recall that :

$$\Rightarrow \tan x = \frac{\sin(x)}{\cos(x)}$$

This is a rational polynomial form, which is defined for $\cos(x) \neq 0$. Now, cos(x) evaluates to zero for certain values of “x”, which appears at a certain interval given by the condition,

$$\cos(x) = 0; \quad x = (2n+1)\frac{\pi}{2}, \quad \text{where} \quad n \in Z$$

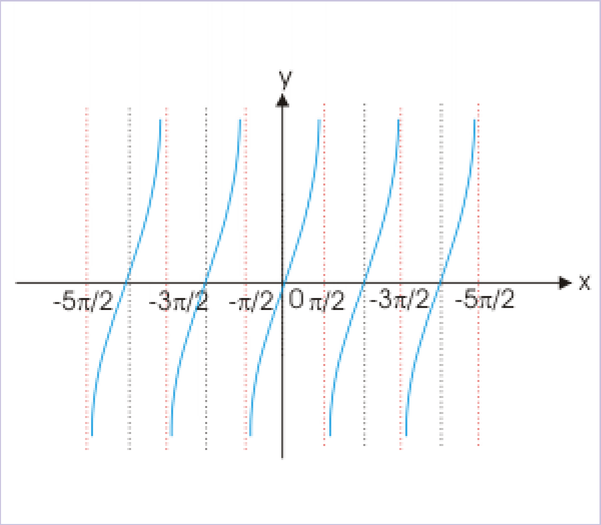
The function cos(x) is zero for $x = \pm\frac{\pi}{2,} \pm \frac{3\pi}{2,} \pm \frac{5\pi}{2,} \ldots$ etc. In other words, the cosine function is zero for all odd multiples of “ $\pi/2$ ”. It means that tangent function is not defined for odd multiples of “ $\pi/2$ ”. Therefore, values of “x” for which cosine is zero need to be excluded from the domain set of real number set “R”. On the other hand, the values of tangent function are extended along the real number line on either side of zero. The range of the function, therefore, is “R”. Hence, domain and range of tangent function are :

$$\text{Domain} = R - \left\{x : x = (2n+1)\frac{\pi}{2}, \quad n \in Z\right\}$$

$$\text{Range} = R$$

The plot of tan(x) .vs. x is shown here.

Tangent function



Graph of tangent function

The period of tanx is π. Multiplication of tangent function by a constant A does not change the range as in the case of sine and cosine function. The plot is always extended on either side of x-axis so that its range is R. Multiplying argument x like tan(kx), however, changes the points where function is not defined. It is now given by :

$$x = (2n+1)\frac{\pi}{2\mathrm{k}}, \quad n \in Z\}$$

Therefore, domain is now modified as :

$$\text{Domain} = R - \left\{x : x = (2n+1)\frac{\pi}{2\mathrm{k}}, \quad n \in Z\right\}$$

Cosecant function

For a real number “x”, there is a cosecant function defined as :

$$f(x) = \operatorname{cosec}(x)$$

Again, the function is not defined for all real number “x”. Let us recall that :

$$\Rightarrow \operatorname{cosec} x = \frac{1}{\sin(x)}$$

This is a rational polynomial form, which is defined for $\sin(x) \neq 0$. Now, sin(x) evaluates to zero for values of “x”, which appears at a certain interval given by the condition,

$$\sin(x) = 0; \quad x = n\pi, \quad \text{where} \quad n \in Z$$

This means that sin(x) is zero for $x = 0, \pm\, \pi, \pm 2\pi, \pm 3\pi, \ldots$ etc. In other words, the sine function is zero for all integer multiples of “ π ”. It means that cosecant function is not defined for integral multiples of “ π ”. Therefore, values of “x”, for which sine is zero, need to be excluded from real number set “R” for defining domain of the function.

On the other hand, values of cosecant function are fall in certain intervals. We have seen that values of sine function is between “-1” and “1”, including end points. Reciprocal of these values are either lesser than “-1” or greater than “1”. Symbolically,

$$\operatorname{cosec} x \leq -1$$

$$\operatorname{cosec} x \geq 1$$

Combining two intervals, using modulus function :

$$|\operatorname{cosec} x| \geq 1$$

The combined interval of the cosecant function, therefore, is :

$$(-\infty, -1] \cup (1,\infty]$$

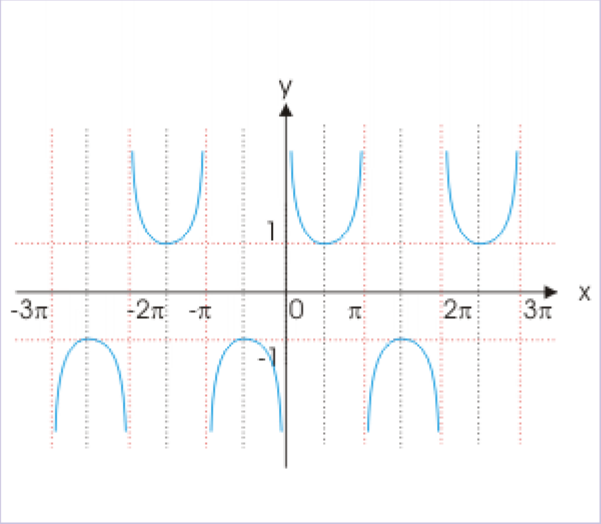
Hence, domain and range of cosecant function are :

$$\text{Domain} = R - \{x : x = n\pi, \quad n \in Z\}$$

$$\text{Range} = (-\infty, -1] \cup (1,\infty]$$

The plot of cosecant(x) .vs. x is shown here.

Cosecant function



Graph of cosecant function

The period of cosecx is 2π. Important to note here is that function is not defined even within a periodic segment. Since cosec(-x) = -cosecx, we conclude that cosecant function is odd function in each of periodic segment. Multiplication of cosecant function by a constant A changes the range as plot lies on or beyond -A or A. The range is modified as :

$$\text{Range} = (-\infty, -\text{A}] \cup (\text{A},\infty]$$

Multiplying argument x like cosec(kx), however, changes the points where function is not defined. It is now given by :

$$x = \frac{n\pi}{k}, \quad n \in Z$$

Therefore, domain is now modified as :

$$\text{Domain} = R - \left\{x : x = \frac{n\pi}{k}, \quad n \in Z\right\}$$

Secant function

For a real number “x”, there is a secant function defined as :

$$f(x) = \sec(x)$$

Again, the function is not defined for all real number “x”. Let us recall that :

$$\sec x = \frac{1}{\cos(x)}$$

This is a rational polynomial form, which is defined for $\cos(x) \neq 0$. Now, cos(x) evaluates to zero for values of “x”, which appears at a certain interval given by the condition,

$$\cos(x) = 0; \quad x = (2n+1)\frac{\pi}{2}, \quad \text{where} \quad n \in Z$$

The function cos(x) is zero for $x = \pm\frac{\pi}{2,} \pm \frac{3\pi}{2,} \pm \frac{5\pi}{2,} \ldots$ etc. In other words, the cosine function is zero for all odd multiples of “ $\pi/2$ ”. It means that secant function is not defined for odd multiples of “ $\pi/2$ ”. These values of “x”, for which cosine is zero, need to be excluded from real number set “R”.

The values of secant function are bounded by certain intervals. We have seen that values of cosine function is between “-1” and “1”, including end points. Just like the case of cosecant function, the range of secant function is :

$$|\sec x| \geq 1$$

or

$$(-\infty, -1] \cup (1,\infty]$$

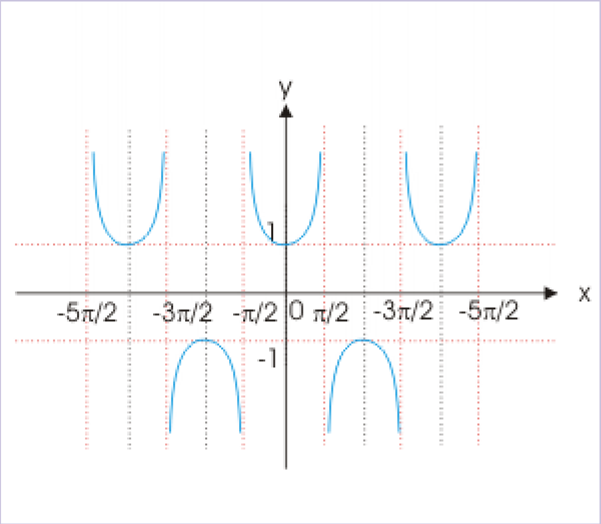
Hence, domain and range of secant function are :

$$\text{Domain} = R - \{x : x = (2n+1)\pi/2, \quad n \in Z\}$$

$$\text{Range} = (-\infty, -1] \cup (1,\infty]$$

The plot of secant(x) .vs. x is shown here.

Secant function



Graph of secant function

The period of secx is 2π. Important to note here is that function is not defined even within a periodic segment. Since sec(-x) = secx, we conclude that secant function is even function in each of periodic segment. Multiplication of secant function by a constant A changes the range plot lies on or beyond -A or A. The range is modified as :

$$\text{Range} = (-\infty, -\text{A}] \cup (\text{A},\infty]$$

Multiplying argument x like sec(kx), however, changes the points where function is not defined. It is now given by :

$$x = (2n+1)\pi/2\text{k}, \quad n \in Z$$

Therefore, domain is now modified as :

$$\text{Domain} = R - \{x : x = (2n+1)\pi/2\text{k}, \quad n \in Z\}$$

Cotangent function

For a real number “x”, there is a cotangent function defined as :

$$f(x) = \cot(x)$$

The function is not defined for all real number “x”. Let us recall that :

$$\Rightarrow \cot x = \frac{\cos(x)}{\sin(x)}$$

This is a rational polynomial form, which is defined for $\sin(x) \neq 0$. Now, sin(x) evaluates to values of “x”, which appears at a certain interval given by the condition,

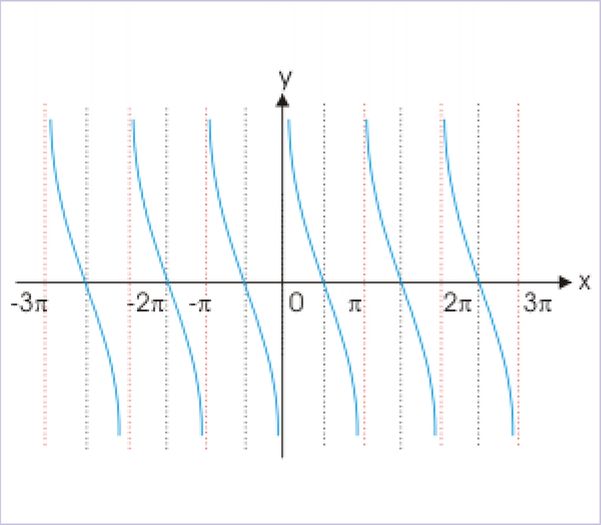
$$\sin(x) = 0; \quad x = n\pi, \quad \text{where} \quad n \in Z$$

This means that sin(x) is zero for $x = 0, \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$ etc. In other words, the sine function is zero for all integral multiples of “ π ”. It means that cotangent function is not defined for integral multiples of “ π ”. Values of “x”, for which sine is zero, need to be excluded from real number set “R”. On the other hand, the values of cotangent function are extended along the real number line on either side of zero. The range of the function, therefore, is “R”. Hence, domain and range of cotangent function are :

$$\text{Domain} = R - \{x : x = n\pi, \quad n \in Z\}$$

$$\text{Range} = R$$

The plot of cot(x) .vs. x is shown here.
Cotangent function



Graph of cotangent function

The period of cotx is π. Since cot(-x) = -cotx, we conclude that cotangennt function is odd function in each of periodic segment. Multiplication of cotangennt function by a constant A does not change the range plot extends either side of x-axis. Multiplying argument x like cot(kx), however, changes the points where function is not defined. It is now given by :

$$x = \frac{n\pi}{k}, \quad n \in Z$$

Therefore, domain is now modified as :

$$\text{Domain} = R - \left\{ x : x = \frac{n\pi}{k}, \quad n \in Z \right\}$$

Examples

Example:
Problem : Find range of the function :

$$f(x) = \frac{1}{2 - \sin 2x}$$

Solution : The denominator of given function is non-negative as value of sin2x can not exceed 1. We know range of sin2x. We shall build up expression from this basic trigonometric function to determine range of the given function. Here,

$$-1 \le \sin 2x \le 1$$

Multiplying with -1 to change sign of sine function, we have (note the change in inequality sign) :

$$1 \ge -\sin 2x \ge -1$$

$$2 + 1 \ge 2 - \sin 2x \ge 2 - 1$$

$$3 \ge 2 - \sin 2x \ge 1$$

We need to take reciprocal of each term in the equality to obtain required function form (note the change in inequality sign),

$$\frac{1}{3} \le \frac{1}{2 - \sin 2x} \le 1$$

$$\frac{1}{3} \le f(x) \le 1$$

$$\text{Range} = \left(\frac{1}{3,} 1\right)$$

This is a unique method to determine range by building up a function from a basic function along with change in the interval of values. We need to be careful that such building up of function does not introduce condition in which function becomes indeterminate. Further, we can find range conventionally by solving function for x in terms of y. We have not considered this method here as solution for x involves inverse trigonometric function. We shall, however, revisit this problem subsequent to the study of inverse trigonometric function.

Example:

Problem : Find range of function :

$$f(x) = 2\sin\sqrt{\left(\frac{\pi^2}{4} - x^2\right)}$$

Solution : Before we attempt to find range, we need to find domain of the function so that we can determine interval of function values. We know that expression within square root is non-negative. Also, expression is a quadratic function. Analyzing this quadratic function, domain of quadratic function is found as [-π/2, π/2]. Coefficient of squared term is negative. Hence, its graph opens down and maximum value of quadratic function is :

$$y_{\max} = -\frac{D}{4a} = -\frac{0 - \left(4X - 1X\frac{\pi}{2}\right)}{4X - 1} = \frac{\pi}{2}$$

Since expression is non-negative within square root, minimum value of function is 0 (see figure). Now, sine function is an increasing function in [0, π/2] as is evident from its graph. Thus sine function assumes values in the interval [sin0, sin π/2] i.e. [0,1]. Sine function, however, has a coefficient of 2 . As such, range of given function is [0,2].

Exercise

Exercise:

Problem: **Problem :** Find the domain of the function given by :

$$f(x) = \sqrt{\cos(\sin x)}$$

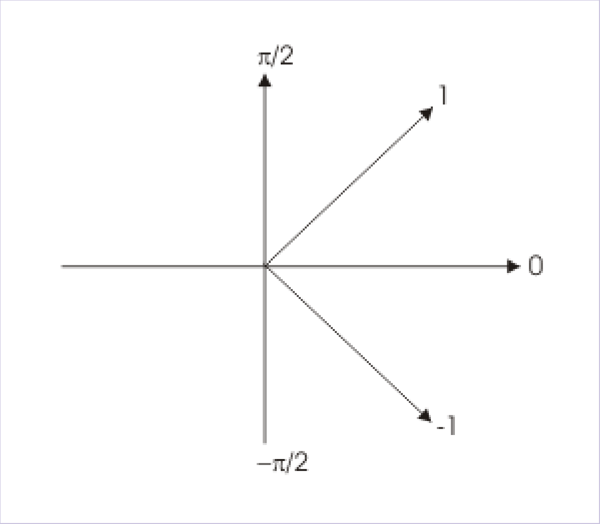
Solution:

Solution : The argument (input) to cosine function is sine function. The expression within square root is non-negative. It means that :

$$\Rightarrow \cos(\sin x) \geq 0$$

We know that cosine function is positive in first and fourth quadrants. It means that the argument of the cosine function should be between -π/2 and π/2. Therefore, we need to see whether the value of "sinx" falls within this interval or not? The value of sine function, on the other hand, lies in the interval [-1,1]. This is indeed (as shown in the figure below) within the required interval for cosine function to be non-negative as 1 < π/2 and -1 > -π/2.

Domain of cosine function



The range of sine function falls within domain of cosine function

Now, we know that sine function is real for all real values of "x". Hence, domain of the given function is :

$$\text{Domain} = (-\infty, \infty)$$

Exercise:

Problem: Problem : Find the domain of the function given by :

$$f(x) = \cos \frac{2\pi}{[x-1]}$$

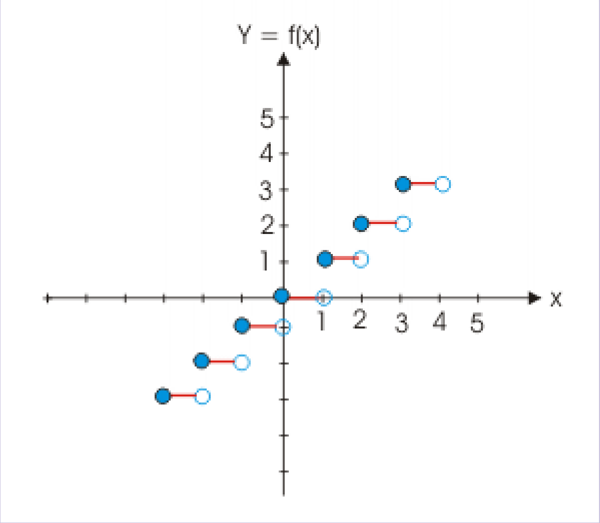
Solution:

Solution : The cosine function is valid for all real values of its argument. The argument, however, is in rational form, requiring that denominator is not zero. Hence,

$$[x-1] \neq 0$$

We can easily evaluate this inequality knowing the fact that greatest integer function is zero in the interval 0≤x<1. This is also substantiated by the graph of greatest integer function as shown here. Now, applying to the greatest integer function of the denominator, the interval in which greatest integer function is equal to zero is :

Greatest integer function



The greatest integer function [x] returns zero in the interval 0≤x<1.

$$0 \le x-1<1 \quad \Rightarrow 1 \le x < 2$$

Hence, domain of the given function is :

$$\text{Domain} = R - [1{,}2)$$

Exercise:

Problem:

Check the validity of the composition and find domain of the function given by :

$$f(x) = \sin\left[\log e \left\{ \frac{\sqrt{(9-x^2)}}{1-x} \right\}\right]$$

Solution:

We know that range of logarithmic function is “R”. Here, logarithmic function itself is the argument of sine function. This means that argument of sine function is "R". This meets the requirement of a sine function, whose domain is "R". Hence, composition of function as given in the question is a valid composition.

In order to find, domain of the function, we need to find values of “x” for which argument of logarithmic function is a positive real number. In the nutshell, we need to evaluate following inequality :

$$\Rightarrow \frac{\sqrt{(9-x^2)}}{1-x} > 0$$

Here, we see that numerator is a square root of a polynomial and is, therefore, positive. Now, evaluating of the polynomial in the numerator for positive real number, we have :

$$\Rightarrow 9 - x^2 > 0$$

$$\Rightarrow x^2 - 9 < 0$$

$$\Rightarrow (x+3)(x-3) < 0$$

$$\Rightarrow -3 < x < 3$$

Since numerator is positive, the denominator needs to be positive so that total rational polynomial is positive. Evaluating inequality relation for the polynomial in the denominator for positive real number :

$$\Rightarrow 1 - x > 0$$

$$\Rightarrow x < 1$$

Now, we know that the domain of the quotient is $D_1 \cap D_2$. Hence,

$$\text{Domain} = (-3 < x < 3) \cap (x < 1)$$

$$\Rightarrow \text{Domain of "f"} = -3 < x < 1$$

Exercise:

Problem: Find range of function :

$$f(x) = \cos\sqrt{\left(x^2 - \frac{\pi^2}{9}\right)}$$

Solution:

Range is [-1,1]

Exercise:

Problem: Find range of function :

$$f(x) = \tan\sqrt{\left(\frac{\pi^2}{9} - x^2\right)}$$

Solution:

Range is [0,√3]

Trigonometric values, equations and identities

In this module, we discuss trigonometric values and angles. In particular, we shall learn about two very useful algorithms which help us to find (i) value of trigonometric function when angle is given and (ii) angles when value of trigonometric function is given. In addition, we shall go through various trigonometric equations and identities. We are expected to be already familiar with them. For this reason, solutions of equations and identities are presented here without deduction and are included for reference purpose.

Values of trigonometric function

It is sufficient to know values of trigonometric functions for angles in first quarter. These angles are called acute angles (angle value less than π/2). Here, we develop algorithm, which converts angles in other quadrants in terms of acute angles. Basic idea is that angles can be expressed in terms of combination of acute angle and reference angles like 0, π/2, π and 2π. These angles demark quadrants. Using certain procedure, we can find value of trigonometric function of any angle provided we know the trigonometric value of corresponding acute angle. For the sake of convenience, we shall concentrate on acute angles π/6, π/4 and π/3, whose trigonometric function values are known to us. We follow an algorithm to determine trigonometric values as given here :

1 : Express given angle as sum or difference of acute angle and reference angles 0, π/2, π and 2π.

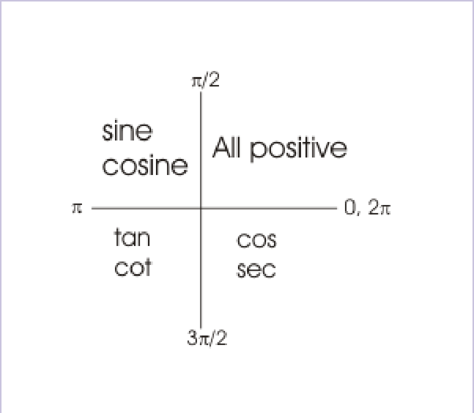
2 : Write trigonometric function of sum or difference as trigonometric function of acute angle. A trigonometric sum/difference combination of angles involving angles of 0, π and 2π does not change the function. However, a combination involving π/2 changes function from sine to cosine and vice-versa, tangent to cotangent and vice-versa and cosecant to secant and vice-versa.

3 : Apply sign before trigonometric function determined as above in accordance with the sign rule of trigonometric function.

$$f(r+a) = (+ \quad or \quad -)g(a)$$

where "f" and "g" denote trigonometric functions, "r" denotes reference angles like 0, π/2, π and 2π and "a" denotes acute angle.

Trigonometric sign diagram



Signs of six trigonometric
functions in different quadrants.

Let us consider an angle 7π/6. We are required to find sine and cotangent values of this angle. Here, we see that 7π/6 is greater than π. Hence, it is equal to π plus some acute angle, say, x.

$$\pi + x = \frac{7\pi}{6}$$

$$\Rightarrow x = \frac{7\pi}{6} - x = \frac{\pi}{6}$$

$$\Rightarrow \sin\frac{7\pi}{6} = \sin\left(\pi + \frac{\pi}{6}\right)$$

Since combination involves angle π, the sine of given angle retains the trigonometric function form. However, angle 7π/6 falls in third quadrant, in which sine is negative. Thus,

$$\Rightarrow \sin\frac{7\pi}{6} = \sin\left(\pi + \frac{\pi}{6}\right) = -\sin\frac{\pi}{6} = -\frac{\sqrt{3}}{2}$$

Similarly,

$$\Rightarrow \cot\frac{7\pi}{6} = \cot\left(\pi + \frac{\pi}{6}\right) = \cot\frac{\pi}{6} = \frac{1}{\sqrt{3}}$$

This method is very helpful to determine value of trigonometric function provided we know the value of trigonometric function of corresponding acute angle resulting from combination involving angles 0, π/2, π and 2π. Here, we shall work out few standard identities involving combination of angles with reference angles. We need not remember these identities. Rather, we should rely on the procedure discussed here as all of these can be derived on spot easily.

Reflection in 0

There is no change in function form. Function takes sign in accordance with sign rule.

$$\sin(0 - x) = -\sin x; \quad \cos(0 - x) = \cos x; \quad \tan(0 - x) = -\tan x;$$

$$\operatorname{cosec}(0 - x) = -\operatorname{cosec} x; \quad \sec(0 - x) = \sec x; \quad \cot(0 - x) = -\cot x$$

Reflection in π/2

Reflection in π/2 is also known as co-function identities. Functions are called co-functions when their compliments have same value. As such, sine and cosine are co-functions. In this case, there is change in function form as combination of angle involves π/2. Function takes sign in accordance with sign rule.

$$\sin\left(\frac{\pi}{2}-x\right)=\cos x;\quad \cos\left(\frac{\pi}{2}-x\right)=\sin x;\quad \tan\left(\frac{\pi}{2}-x\right)=\cot x;$$

$$\operatorname{cosec}\left(\frac{\pi}{2}-x\right)=\sec x;\quad \sec\left(\frac{\pi}{2}-x\right)=\operatorname{cosec} x;\quad \cot\left(\frac{\pi}{2}-x\right)=\tan x$$

Reflection in π

In this case, there is no change in function form. Function takes sign in accordance with sign rule.

$$\sin(\pi-x)=\sin x;\quad \cos(\pi-x)=-\cos x;\quad \tan(\pi-x)=-\tan x;$$

$$\operatorname{cosec}(\pi-x)=\operatorname{cosec} x;\quad \sec(\pi-x)=-\sec x;\quad \cot(\pi-x)=-\cot x$$

Shift by π/2

Shift refers to horizontal shift of graph. We shall explore this aspect of trigonometric function in detail in a separate module. From transformation point of view, there is change in function form as combination of angle involves π/2. Function takes sign in accordance with sign rule.

$$\sin\left(\frac{\pi}{2}+x\right)=\cos x;\quad \cos\left(\frac{\pi}{2}+x\right)=-\sin x;\quad \tan\left(\frac{\pi}{2}+x\right)=-\cot x;$$

$$\operatorname{cosec}\left(\frac{\pi}{2}+x\right)=\sec x;\quad \sec\left(\frac{\pi}{2}+x\right)=-\operatorname{cosec} x;\quad \cot\left(\frac{\pi}{2}+x\right)=-\tan x$$

Shift by π

In this case, there is no change in function form. Function takes sign in accordance with sign rule.

$$\sin(\pi+x)=-\sin x;\quad \cos(\pi+x)=-\cos x;\quad \tan(\pi+x)=\tan x;$$

$$\operatorname{cosec}(\pi+x)=-\operatorname{cosec} x;\quad \sec(\pi+x)=-\sec x;\quad \cot(\pi+x)=\cot x$$

Shift by 2π

In this case, there is no change in function form. Function takes sign in accordance with sign rule.

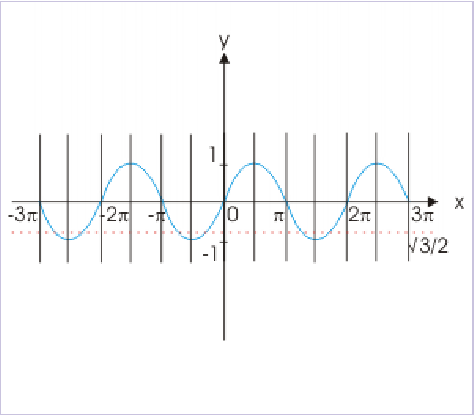
$$\sin(2\pi + x) = \sin x; \quad \cos(2\pi + x) = \cos x; \quad \tan(2\pi + x) = \tan x;$$

$$\operatorname{cosec}(2\pi + x) = \operatorname{cosec} x; \quad \sec(2\pi + x) = \sec x; \quad \cot(2\pi + x) = \cot x$$

Finding angles

Trigonometric functions are many-one relation. We are required to find angles corresponding to a given trigonometric value. For example, what are angles corresponding to sine value of -√3/2. In other words, we need to find angles whose sine evaluates to this value. Note that these values corresponds to intersection of parallel line y=-√3/2 with the graph of sine curve.

Graph of sine function



Intersection of sine function with parallel value line.

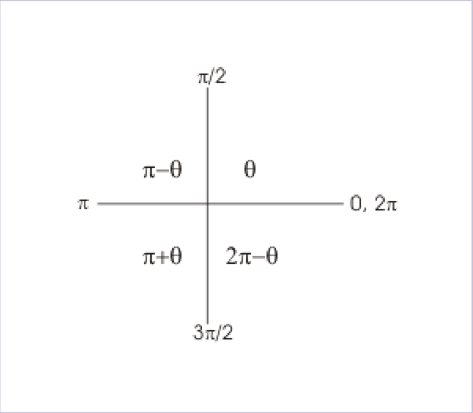
For the time being, let us concentrate the interval [0,2π], which corresponds to one cycle of four quadrants. We follow an algorithm as given here to find angles in this interval :

1 : Consider only numerical magnitude of the given value. Find acute angle whose trigonometric function value corresponds to the numerical magnitude of the given value.

2 : Use sign rule and identify quadrants in which trigonometric function has the sign that of given value.

3 : Use value diagram and determine the angles as required.

Trigonometric value diagram



Angles whose trigonometric function values are same in different quadrants(to be used in conjunction with sign diagram).

To see the working of the algorithm, let us consider sinx = -√3/2. Considering only the magnitude of numerical value, we have :

$$\sin\theta = \frac{\sqrt{3}}{2} = \sin\frac{\pi}{3}$$

Thus, required acute angle is π/3. Now, sine function is negative in third and fourth quadrants. Looking at the value diagram, the angle in third quadrant is :

$$\Rightarrow x = \pi + \theta = \pi + \frac{\pi}{3} = \frac{4\pi}{3}$$

Similarly, angle in fourth quadrant is :

$$\Rightarrow x = 2\pi - \theta = 2\pi - \frac{\pi}{3} = \frac{5\pi}{3}$$

Example:

Problem : Find angles in [0,2π], if

$$\cot x = \frac{1}{\sqrt{3}}$$

Solution : Considering only the magnitude of numerical value, we have :

$$\Rightarrow \cot\theta = \frac{1}{\sqrt{3}} = \cot\frac{\pi}{3}$$

Thus, required acute angle is π/3. Now, cotangent function is positive in first and third quadrants. Looking at the value diagram, the angle in third quadrant is :

$$\Rightarrow x = \pi + \theta = \pi + \frac{\pi}{3} = \frac{4\pi}{3}$$

Hence angles are π/3 and 4π/3.

Negative angles

When we consider angle as a real number entity, we need to express angles as negative angles as well. The corresponding negative angle (y) is obtained as :

$$y = x - 2\pi$$

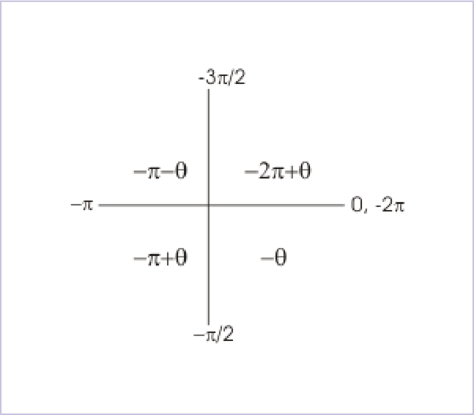
Thus, negative angles corresponding to 4π/3 and 5π/3 are :

$$\Rightarrow y = \frac{4\pi}{3} - 2\pi = -\frac{2\pi}{3}$$

$$\Rightarrow y = \frac{5\pi}{3} - 2\pi = -\frac{\pi}{3}$$

We can also find negative angle values using a separate negative value diagram (see figure). We draw negative value diagram by demarking quadrants with corresponding angles and writing angle values for negative values. We deduct “2π” from the relation for positive value diagram.

Trigonometric value diagram



Trigonometric value diagram
for negative angles

Let us consider sinx = -√3/2 again. The acute angle in first quadrant is π/3. Sine is negative in third and fourth quadrants. The angles in these quadrants are :

$$y = -\pi + \theta = -\pi + \frac{\pi}{3} = -\frac{2\pi}{3}$$

$$y = -\theta = -\frac{\pi}{3}$$

Trigonometric equations

Zeroes of sine and cosine functions

Trigonometric equations are formed by equating trigonometric functions to zero. The solutions of these equations are :

1 : $\sin x = 0 \quad \Rightarrow x = n\pi; n \in Z$

2 : $\cos x = 0 \quad \Rightarrow x = (2n+1)\frac{\pi}{2}; n \in Z$

Definition of other trigonometric functions

We define other trigonometric functions in the light of zeroes of sine and cosine as listed above :

$$\tan x = \frac{\sin x}{\cos x}; x \neq (2n+1)\frac{\pi}{2}; n \in Z$$

$$\cot x = \frac{\cos x}{\sin x}; x \neq n\pi; n \in Z$$

$$\operatorname{cosec} x = \frac{1}{\sin x}; x \neq n\pi; n \in Z$$

$$\sec x = \frac{1}{\cos x}; x \neq (2n+1)\frac{\pi}{2}; n \in Z$$

Trigonometric equations

Trigonometric function can be used to any other values as well. Solutions of such equations are given here without deduction for reference purpose. Solutions of three equations involving sine, cosine and tangent functions are listed here :

1. Sine equation

$$\sin x = a = \sin y$$

$$x = n\pi + (-1)^n y; n \in Z$$

2. Cosine equation

$$\cos x = a = \cos y$$

$$x = 2n\pi \pm y; n \in Z$$

3. Tangent equation

$$\tan x = a = \tan y$$

$$x = n\pi + y; n \in Z$$

In order to understand the working with trigonometric equation, let us consider an equation :

$$\sin x = -\frac{\sqrt{3}}{2}$$

As worked out earlier, -√3/2 is sine value of two angles in the interval [0, π]. Important question here is to know which angle should be used in the solution set. Here,

$$\Rightarrow \sin\frac{4\pi}{3} = \sin\frac{5\pi}{3} = -\frac{\sqrt{3}}{2}$$

We can write general solution using either of two values.

$$x = n\pi + (-1)^n \frac{4\pi}{3}; n \in Z$$

$$\Rightarrow x = n\pi + (-1)^n \frac{5\pi}{3}; n \in Z$$

The solution sets appear to be different, but are same on expansion. Conventionally, however, we use the smaller of two angles which lie in the interval [0, π]. In order to check that two series are indeed same, let us expand series from n=-4 to n=4,

For $\Rightarrow x = n\pi + (-1)^n \frac{4\pi}{3}; n \in Z$

$$-4\pi + \frac{4\pi}{3} = -\frac{8\pi}{3}, -3\pi - \frac{4\pi}{3} = -\frac{13\pi}{3}, -2\pi + \frac{4\pi}{3} = -\frac{2\pi}{3}, -\pi - \frac{4\pi}{3} = -\frac{7\pi}{3},$$

$$0 + 4\pi/3 = \frac{4\pi}{3}, \pi - \frac{4\pi}{3} = -\frac{\pi}{3}, 2\pi + \frac{4\pi}{3} = \frac{10\pi}{3}, 3\pi - \frac{4\pi}{3} = \frac{5\pi}{3}, 4\pi + \frac{4\pi}{3} = \frac{16\pi}{3}$$

Arranging in increasing order :

$$-\frac{13\pi}{3}, -\frac{8\pi}{3}, -\frac{7\pi}{3}, -\frac{2\pi}{3}, -\frac{\pi}{3}, \frac{4\pi}{3}, \frac{5\pi}{3}, \frac{10\pi}{3}, \frac{16\pi}{3}$$

For $\Rightarrow x = n\pi + (-1)^n \frac{5\pi}{3}; n \in Z$

$$-4\pi + \frac{5\pi}{3} = -\frac{7\pi}{3}, -3\pi - \frac{5\pi}{3} = -\frac{14\pi}{3}, -2\pi + \frac{5\pi}{3} = -\frac{\pi}{3}, -\pi - \frac{5\pi}{3} = -\frac{8\pi}{3},$$

$$0+\frac{5\pi}{3}=\frac{5\pi}{3},\pi-\frac{5\pi}{3}=-\frac{2\pi}{3},2\pi+\frac{5\pi}{3}=\frac{11\pi}{3},3\pi-\frac{5\pi}{3}=\frac{4\pi}{3},4\pi+\frac{5\pi}{3}=\frac{17\pi}{3}$$

Arranging in increasing order :

$$-\frac{14\pi}{3},-\frac{8\pi}{3},-\frac{7\pi}{3},-\frac{2\pi}{3},-\frac{\pi}{3},\frac{4\pi}{3},\frac{5\pi}{3},\frac{11\pi}{3},\frac{17\pi}{3}$$

We see that there are common terms. There are, however, certain terms which do not appear in other series. We can though find those missing terms by evaluating some more values. For example, if we put n = 6 in the second series, then we get the missing term -13π/3. Also, putting n=5,7, we get 10π/3 and 16π/3. Thus, all missing terms in second series are obtained. Similarly, we can compute few more values in first series to find missing terms. We, therefore, conclude that both these series are equal.

Example:

Problem : Find solution of equation :

$$2\cos^2 x+3\sin x=0$$

Solution : Our objective here is to covert equation to linear form. Here, we can not convert sine term to cosine term, but we can convert $\cos^2 x$ in terms of $\sin^2 x$.

$$\Rightarrow 2\left(1-\sin^2 x\right)+3\sin x=0$$

$$\Rightarrow 2-2\sin^2 x+3\sin x=0$$

$$\Rightarrow 2\sin^2 x-3\sin x-2=0$$

It is a quadratic equation in sinx. Factoring, we have :

$$\Rightarrow 2\sin^2 x+\sin x-4\sin x-2=0$$

$$\Rightarrow \sin x(2\sin x+1)-2(2\sin x+1)=0$$

$$\Rightarrow (2\sin x+1)(\sin x-2)=0$$

Either, sinx=-1/2 or sinx = 2. But sinx can not be equal to 2. hence,

$$\Rightarrow \sin x=-\frac{1}{2}=\sin\left(\pi+\frac{\pi}{6}\right)=\sin\left(\frac{7\pi}{6}\right)$$

$$\Rightarrow x=n\pi+(-1)^n\frac{7\pi}{6};\quad n\in Z$$

Note : We shall not work with any other examples here as purpose of this module is only to introduce general concepts of angles, identities and equations. These topics are part of separate detailed study.

Trigonometric identities

Reciprocal identities

Reciprocals are defined for values of x for which trigonometric function in the denominator is not zero.

$$\sin x = \frac{1}{\operatorname{cosec} x}; \quad \cos x = \frac{1}{\sec x}; \quad \tan x = \frac{1}{\cot x};$$

$$\operatorname{cosec} x = \frac{1}{\sin x}; \quad \sec x = \frac{1}{\cos x}; \quad \cot x = \frac{1}{\tan x}$$

Negative angle identities

$$\cos(-x) = \cos x; \quad \sin(-x) = -\sin x; \quad \tan(-x) = -\tan x$$

Pythagorean identities

$$\cos^2 x + \sin^2 x = 1; \quad 1 + \tan^2 x = \sec^2 x; \quad 1 + \cot^2 x = \operatorname{cosec}^2 x$$

Sum/difference identities

$$\sin(x \pm y) = \sin x \cos y \pm \sin y \cos x$$

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$

$\tan(x \pm y) = \tan sx \pm \tan y/1 \mp \tan x \tan y;$ x,y and (x+y) are not odd multiple of π/2

$\cot(x \pm y) = \cot x \cot y \mp 1/\cot y \pm \cot x;$ x,y and (x+y) are not odd multiple of π/2

Double angle identities

$$\sin 2x = 2 \sin x \cos x = \frac{2 \tan x}{1 + \tan^2 x}$$

$$\cos 2x = \cos^2 x - \sin^2 x = 2\cos^2 x - 1 = 1 - 2\sin^2 x = \frac{1 - \tan^2 x}{1 + \tan^2 x}$$

$$\tan 2x = \frac{2 \tan x}{1 - \tan^2 x}$$

$$\cot 2x = \frac{\cot^2 x - 1}{2\cot x}$$

Triple angle identities

$$\sin 3x = 3\sin x - 4\sin^3 x$$

$$\cos 3x = 4\cos^3 x - 3\cos x$$

$$\tan 3x = \frac{3\tan x - \tan^3 x}{1 - 3\tan^2 x}$$

$$\cot 3x = \frac{3\cot x - \cot^3 x}{1 - 3\cot^2 x}$$

Power reduction identities

$$\sin{}^2 x = \frac{1 - \cos 2x}{2}$$

$$\cos{}^2 x = \frac{1 + \cos 2x}{2}$$

$$\sin{}^3 x = \frac{3\sin x - \sin 3x}{4}$$

$$\cos^3 x = \frac{\cos 3x + 3\cos x}{4}$$

Product to sum identities

$$2\sin x\cos y = \sin(x+y) + \sin(x-y)$$

$$2\cos x\sin y = \sin(x+y) - \sin(x-y)$$

$$2\cos x\cos y = \cos(x+y) + \cos(x-y)$$

$$2\sin x\sin y = -\cos(x+y) + \cos(x-y) = \cos(x-y) - \cos(x+y)$$

Sum to product identities

$$\sin x + \sin y = 2\sin\frac{(x+y)}{2}\cos\frac{(x-y)}{2}$$

$$\sin x - \sin y = 2\cos\frac{(x+y)}{2}\sin\frac{(x-y)}{2}$$

$$\cos x + \cos y = 2\cos\frac{(x+y)}{2}\cos\frac{(x-y)}{2}$$

$$\cos x - \cos y = -2\sin\frac{(x+y)}{2}\sin\frac{(x-y)}{2} = 2\sin\frac{(x+y)}{2}\sin\frac{(y-x)}{2}$$

Half angle identities

$$\sin\frac{x}{2} = \pm\sqrt{\left\{\frac{(1-\cos x)}{2}\right\}}$$

$$\cos\frac{x}{2} = \pm\sqrt{\left\{\frac{(1+\cos x)}{2}\right\}}$$

$$\tan\frac{x}{2} = \operatorname{cosec} x - \cot x = \pm\sqrt{\left\{\frac{(1-\cos x)}{(1+\cos x)}\right\}} = \frac{\sin x}{1+\cos x} = \frac{1-\cos x}{\sin x}$$

$$\cot\frac{x}{2} = \operatorname{cosec} x + \cot x = \pm\sqrt{\left\{\frac{(1+\cos x)}{(1-\cos x)}\right\}} = \frac{\sin x}{1-\cos x} = \frac{1+\cos x}{\sin x}$$

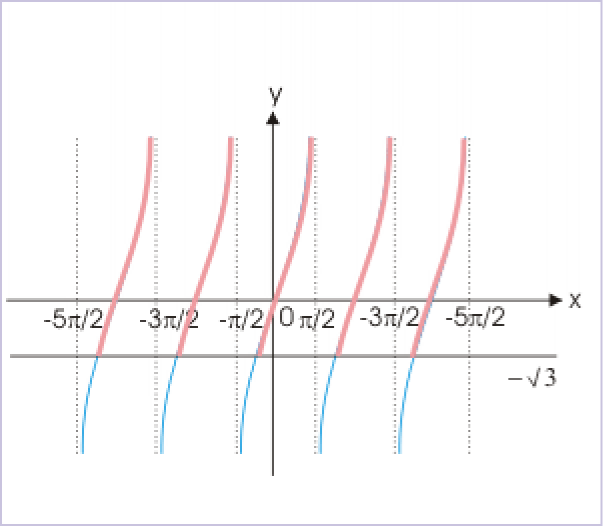
Trigonometric inequalities

Evaluating trigonometric ratios is a direct process in which we make use of known values, trigonometric identities and transformations or even pre-defined trigonometric tables. The evaluation of trigonometric inequalities is somewhat inverse of this process. Consider an inequality :

$$\tan x \geq -\sqrt{3}$$

Clearly, we need to know "x" for which this inequality holds. As pointed out earlier, trigonometric functions are "many-one" relation. The value of "x" satisfying given inequality is not an unique interval, but a series of intervals. Incidentally, however, trigonometric function values repeat after certain "period". So this enables us to define periodic intervals in generic manner for which trigonometric inequality holds.

Trigonometric inequality



Intervals satisfying inequality, involving tangent function.

We observe that line $y = -\sqrt{3}$ intersects tangent graph at multiple points. The sections of plots satisfying the inequality are easily identified on the graph and are shown as dark red line.

Solution of trigonometric inequality

Determination of base or fundamental interval is central to solve trigonometric inequality. The function values in this interval is repeated with a periodicity of trigonometric function. The base interval depends on the nature of trigonometric function and inequality in question. The steps to find solution of trigonometric inequality are :

1 : Convert given inequality to trigonometric equation by replacing inequality sign by equality sign.

2 : Solve resulting equation in the interval [0,2π]. There are two solutions. They are the angle values at which trigonometric function has the value which is being compared in the given inequality.

3 : Convert positive angle greater than π to equivalent negative value to account for the fact that basic interval being repeated may lie on negative side of the origin (cosine, secant and tangent function).

4 : Construct base interval between two values, keeping in mind the given inequality. It is always advantageous to draw a rough intersection of graphs of each side of given inequality.

5 : If function asymptotes (tangent, cotangent, secant and cosecant) within the interval constructed, then basic interval is limited by the angle value at which function asymptotes.

6 : Generalize solution by extending base interval with the period of the trigonometric function.

In order to understand the process, let us solve the inequality given by :

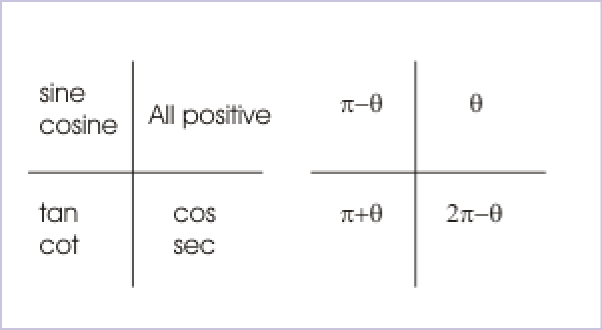
$$\tan x \geq -\sqrt{3}$$

This example has been selected here as it involves consideration of each step as enumerated above for finding solution of inequality. Corresponding trigonometric equation, in this case, is :

$$\tan x = -\sqrt{3}$$

The acute angle is π/3. Further, tangent function is negative in second and fourth quarter (see sign diagram). Using value diagram in conjunction with sign diagram, solution of given equation in [0, 2π] are :

Sign and value diagram



Tangent function is negative in second and fourth quarter .

$$\Rightarrow x = \pi - \theta = \pi - \frac{\pi}{3} = \frac{2\pi}{3}$$

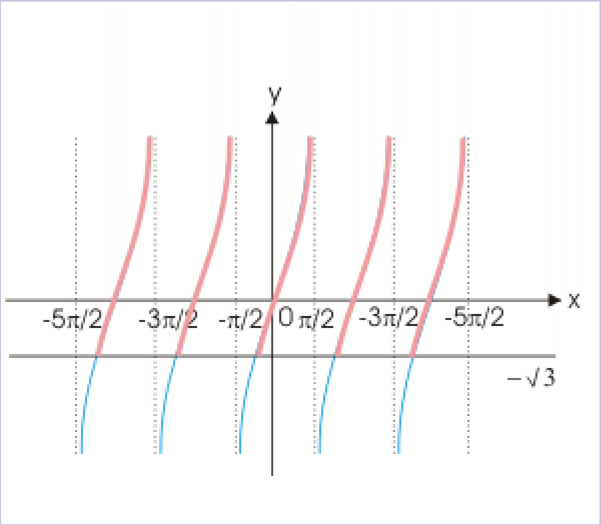
$$\Rightarrow x = 2\pi - \theta = 2\pi - \frac{\pi}{3} = \frac{5\pi}{3}$$

Here, second angle is greater than π. Hence, equivalent negative angle is :

$$\Rightarrow y = \frac{5\pi}{3} - 2\pi = -\frac{\pi}{3}$$

Tangent function, however, is not a continuous function between -π/3 and 2π/3. Tangent values are greater than -√3 for angle greater than -π/3, but value asymptotes to infinity at π/2. This can be verified from the intersection graph.

Trigonometric inequality



Intervals satisfying inequality, involving tangent function.

Thus, basic interval satisfying inequality is :

$$-\frac{\pi}{3} \le x < \frac{\pi}{2}$$

It is also clear that the solution in this interval is repeated with a period of π, which is period of tangent function. Hence, solution of given inequality is :

$$n\pi - \frac{\pi}{3} \le x < n\pi + \frac{\pi}{2}; \quad n \in Z$$

Examples

Example:
Problem : Solve trigonometric inequality given by :

$$\sin x \ge \frac{1}{2}$$

Solution : The solution of the corresponding equal equation is obtained as :

$$\sin x = \frac{1}{2} = \sin \frac{\pi}{6}$$

$$\Rightarrow x = \frac{\pi}{6}$$

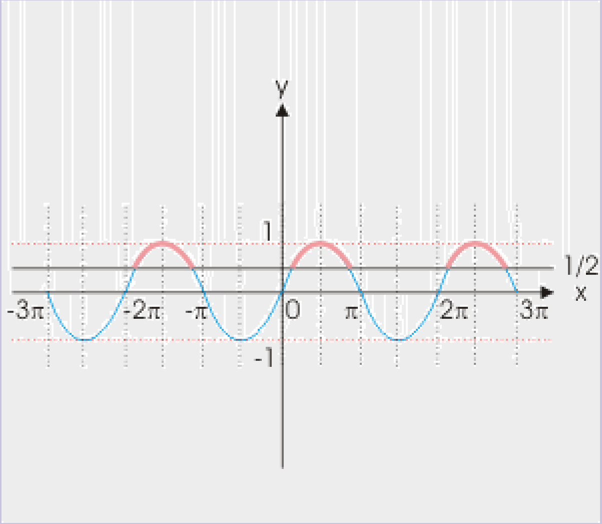
The sine function is positive in first and second quarter. Hence, second angle between “0” and “2π” is :

$$\Rightarrow x = \pi - \theta = \pi - \frac{\pi}{6} = \frac{5\pi}{6}$$

Both angles are less than “π”. Thus, we do not need to convert angle into equivalent negative angle. Further, sine curve is defined for all values of “x”. The base interval, therefore, is :

The valid intervals on sine plot are shown in the figure.

Trigonometric inequality



Intervals satisfying inequality

$$\frac{\pi}{6} \le x \le \frac{5\pi}{6}$$

The periodicity of sine function is “2π”. Hence, we add “2nπ” on either side of the base interval :

$$2n\pi + \frac{\pi}{6} \le x \le 2n\pi + \frac{5\pi}{6}, \quad n \in Z$$

Example:

Problem : Solve trigonometric inequality given by :

$$\sin x > \cos x$$

Solution : In order to solve this inequality, it is required to convert it in terms of inequality of a single trigonometric function.

$$\sin x > \cos x$$

$$\Rightarrow \sin x - \cos x > 0$$

$$\Rightarrow \sin x \cos \frac{\pi}{4} - \cos x \sin \frac{\pi}{4} > 0$$

$$\Rightarrow \sin\left(x - \frac{\pi}{4}\right) > 0$$

Let $y = x - \pi/4$. Then,

$$\sin y > 0$$

Thus, we see that problem finally reduces to solving trigonometric sine inequality. The solution of the corresponding equality is obtained as :

$$\Rightarrow \sin y = 0 = \sin 0$$

$$\Rightarrow y = 0$$

The second angle between “0” and “2π” is “π”. The base interval, therefore, is :

$$0 < y < \pi$$

The periodicity of sine function is “2π”. Hence, we add “2nπ” on either side of the base interval :

$$2n\pi < y < 2n\pi + \pi, \quad n \in Z$$

Now substituting for $y = x - \pi/4$, we have :

$$2n\pi < x - \frac{\pi}{4} < 2n\pi + \pi, \quad n \in Z$$

$$\Rightarrow 2n\pi + \frac{\pi}{4} < x < 2n\pi + \frac{5\pi}{4}, \quad n \in Z$$

Example:

Problem : If domain of a function, “f(x)”, is [0,1], then find the domain of the function given by :

$$f(2\sin x - 1)$$

Solution : The domain of the function is given here. We need to find the domain when argument (input) to the function is a trigonometric expression. The given domain is :

$$0 \le x \le 1$$

Changing argument of the function, the domain becomes :

$$0 \le 2\sin x - 1 \le 1 \quad \Rightarrow 1 \le 2\sin x \le 2 \quad \Rightarrow 1/2 \le \sin x \le 1$$

However, the range of sinx is [-1,1]. It means that the above interval is equivalent to a trigonometric inequality given by :

$$\Rightarrow \sin x \ge \frac{1}{2}$$

The sine function is positive in first and second quadrant. Two values of “x” between “0” and “2π” are :

$$\frac{\pi}{6}, \pi - \frac{\pi}{6}$$

$$\Rightarrow \frac{\pi}{6}, \frac{5\pi}{6}$$

The value of “x” satisfying above equation :

$$2n\pi + \pi/6 <= x <= 2n\pi + 5\pi/6, \quad n \in Z$$

Hence, required domain is :

$$\left[2n\pi + \frac{\pi}{6}, 2n\pi + \frac{5\pi}{6}\right], \quad n \in Z$$

Example:
Problem : Find the domain of the function given by :

$$f(x) = \log_e \frac{1}{\sqrt{[\cos x] - [\sin x]}}$$

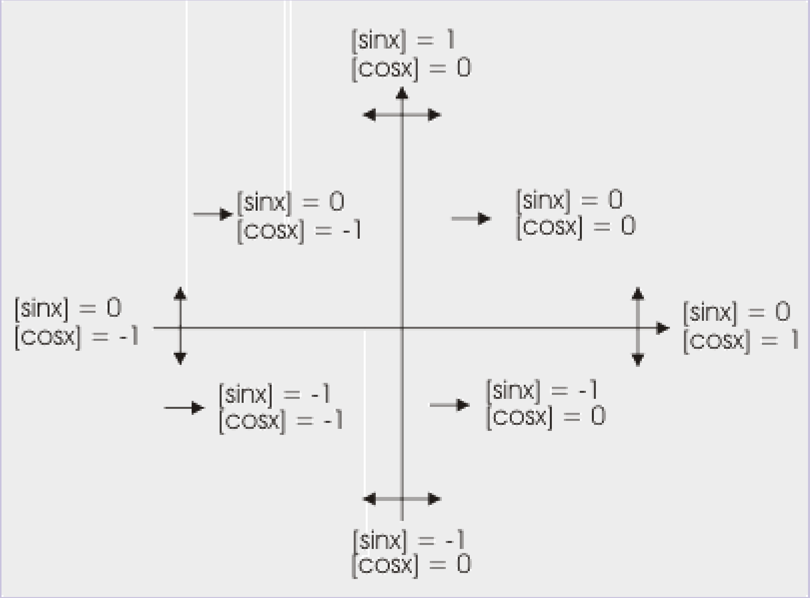
Solution : The function is a logarithmic function, which is valid for all positive values of its argument. Also, the argument of logarithmic function is in rational form, having denominator as a square root. We have to find values of “x” for which the expression within the square root is a positive number. It means that :

$$\Rightarrow [\cos x] - [\sin x] > 0$$

$$\Rightarrow [\cos x] > [\sin x]$$

In order to evaluate this inequality, we determine modulus of two trigonometric functions in four quadrants at all bounding values of angle “x” and also at intermediate angles. The values are shown in the figure :

Modulus of trigonometric functions



The values of modulus of sine and cosine functions in four quadrants are shown.

It is clear that $[\cos x] > [\sin x]$ is true in the fourth quadrant. Hence, domain of the function is :

$$\text{Domain} = -\frac{\pi}{2} \leq x \leq 0 = \left[-\frac{\pi}{2,} 0\right]$$

Exercise

Exercise:

Problem: Solve the inequality :

$$\tan x < 1$$

Solution:

Hints : Here, corresponding trigonometric equation is :

$$\tan x = 1 = \tan\frac{\pi}{4}$$

The other solution in [0,2π] is :

$$x = \pi + \frac{\pi}{4} = \frac{5\pi}{4}$$

Corresponding negative angle :

$$y = \frac{5\pi}{4} - 2\pi = -\frac{3\pi}{4}$$

However, tangent function asymptotes at - π/2. Hence, basic interval is :

$$\left(-\frac{\pi}{2}, \frac{\pi}{4}\right)$$

Further, tangent function has period of π. The general solution is :

$$n\pi - \frac{\pi}{2} < x < n\pi + \frac{\pi}{4}; n \in Z$$

Exercise:

Problem: Find domain of function :

$$y = \log_e \cos x$$

Solution:

Argument of logarithmic function is positive. Hence,

$$\cos x > 0$$

The angles of corresponding trigonometric equation sinx = 0 in the interval [0,2π] are 3π/2 and π/2. In terms of negative angles are -π/2 and π/2. Cosine function is positive in first and fourth quadrant,

$$-\frac{\pi}{2} < x < \frac{\pi}{2}$$

Since cosine values are repeated after "2π", we can write general inequality as :

$$2n\pi - \frac{\pi}{2} < x < 2n\pi + \frac{\pi}{2}; \quad n \in Z$$

$$(4n-1)\frac{\pi}{2} < x < (4n+1)\frac{\pi}{2}; \quad n \in Z$$

The domain of given function, therefore, is :

$$(4n-1)\frac{\pi}{2} < x < (4n+1)\frac{\pi}{2}; \quad n \in Z$$

Exercise:

Problem: Find domain of function :

$$\mathrm{f(x)} = \sqrt{(\sin x - 1)}$$

Solution:

Expression under square root is non-negative. Hence,

$$\Rightarrow \sin x - 1 \geq 0$$

$$\Rightarrow \sin x \geq 1$$

But value of sine function can not be greater than 1. Thus, we need to only evaluate equation sinx=1 to find the domain. Here,

$$\Rightarrow \sin x = 1 = \sin\frac{\pi}{2}$$

The domain of given function, therefore, is :

$$x = \{x : n\pi + (-1)^n \pi/2; \quad n \in Z\}$$

Exercise:

Problem: Find domain of function :

$$f(x) = \frac{1}{\sqrt{(\sin x)}} + \sqrt[3]{}(\sin x)$$

Solution:

We treat this function as addition of two individual functions. Expression under square root in the first function is non-negative. However, radical appears in denominator. As such,

$$\Rightarrow \sin x > 0$$

The angles of corresponding trigonometric equation sinx = 0 in the interval [0,2π] are 0 and π. Since sine function is positive in first and second quadrant,

$$0 < x < \pi$$

Since sine values are repeated after “2π”, we can write general inequality as :

$$0 + 2n\pi < x < 2n\pi + \pi; \quad n \in Z$$

On the other hand, domain of second function is real number set R. Note that it is not even radical. Now, domain of addition of two functions is intersection of individual domains. Hence, domain of given function is :

$$2n\pi < x < (2n+1)\pi; \quad n \in Z$$

Exercise:

Problem:

A function f(x) is defined in [0,1]. Determine range of function definition f(sinx).

Solution:

The domain of f(x) is [0,1]. Here, argument of function changes from “x” to “sinx”. It changes the input value to function for x, but values in themselves lie within the domain interval of the function. It means that :

$$0 \le \sin x \le 1$$

$$\Rightarrow \sin 0 \le \sin x \le \sin \frac{\pi}{2}$$

$$\Rightarrow 0 \le x \le \pi/2$$

This is now a problem of solving two inequalities. The first inequality is : sinx≥0

As solved earlier, the solution is :

$$2n\pi < x < (2n+1)\pi; \quad n \in Z$$

The second inequality is :

$$\sin x \le 1$$

Since value of sinx can not exceed value of 1, we conclude that values of x which satisfies inequality sinx≥0 also satisfies the inequality sinx ≤1. Hence, domain of given function is :

$$2n\pi < x < (2n+1)\pi; \quad n \in Z$$

Exercise:

Problem:

A function f(x) is defined in [0,1]. Determine range of function definition f(tanx).

Solution:

The domain of f(x) is [0,1]. Here, argument of function changes from "x" to "tanx". It changes the input value to function for x, but values in themselves lie within the domain interval of the function. It means that :

$$0 \le \tan x \le 1$$

$$\Rightarrow \tan 0 \le \tan x \le \tan \frac{\pi}{4}$$

$$\Rightarrow 0 \le x \le \frac{\pi}{4}$$

$$0 + n\pi \le x \le n\pi + \frac{\pi}{4}$$

Since tangent values are repeated after "π", we can write general inequality as :

$$n\pi \le x \le n\pi + \frac{\pi}{4} \quad n \in Z$$

The domain of given function, therefore, is :

$$n\pi \le x \le n\pi + \frac{\pi}{4} \quad n \in Z$$

Inverse trigonometric functions

Inverse trigonometric function returns an angle corresponding to a real number, following certain rule. They are inverse functions corresponding to trigonometric functions. The inverse function of sine , for example, is defined as :

$$f(x) = \sin^{-1} x; \quad x \in \left[-1, 1\right]$$

where “x” is a real number, "f(x)" is the angle. Clearly, "f(x)" is the angle, whose sine is “x”. Symbolically,

$$\Rightarrow \sin\left\{f(x)\right\} = \sin\left\{\sin^{-1}(x)\right\} = x$$

In the representation of inverse function, we should treat “-1” as symbol – not as power. In particular,

$$\sin^{-1}(x) \neq \frac{1}{\sin x}$$

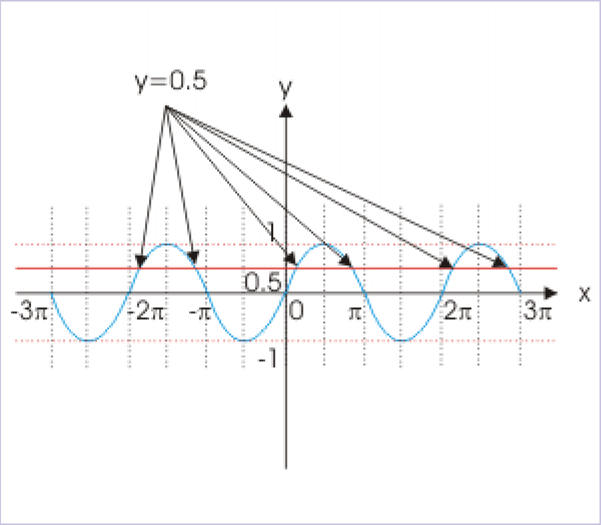
Inverse trigonometric functions are also called arc functions. This is an alternative notation. The corresponding functions are arcsine, arccosine, arctangent etc. For example,

$$\Rightarrow f(x) = \sin^{-1}(x) = \arcsin(x)$$

Nature of trigonometric functions

Trigonometric functions are many-one relations. The trigonometric ratio of different angles evaluate to same value. If we draw a line parallel to x-axis such that 0 < y < 1, then it intersects sine plot for multiple times – ,in fact, infinite times. It follows, then, that we can associate many angles to the same sine value. The trigonometric functions are, therefore, not an injection and hence not a bijection. As such, we can not define an inverse of trigonometric function in the first place! We shall see that we need to redefine trigonometric functions in order to make them invertible.

sine function



many-one relation

In order to define, an inverse function, we require to have one-one relation in both directions between domain and range. The function needs to be a bijection. It emerges that we need to shorten the domain of trigonometric functions such that a distinct angle corresponds to a distinct real number. Similarly, a distinct real number corresponds to a distinct angle.

We can identify many such shortened intervals for a particular trigonometric function. For example, the shortened domain of sine function can be any one of the intervals defined by :

$$\left[(2n-1)\frac{\pi}{2},(2n+1)\frac{\pi}{2}\right], \quad n \in Z$$

The domain corresponding to n = 0 yields principal domain given by :

$$\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$$

The nature of trigonometric functions is periodic. Same values repeat after certain interval. Here, our main task is to identify an interval of “x” such that all possible values of a trigonometric function are included once. This will ensure one-one relation in both directions between domain and range of the function. This interval is easily visible on graphs of the corresponding trigonometric function.

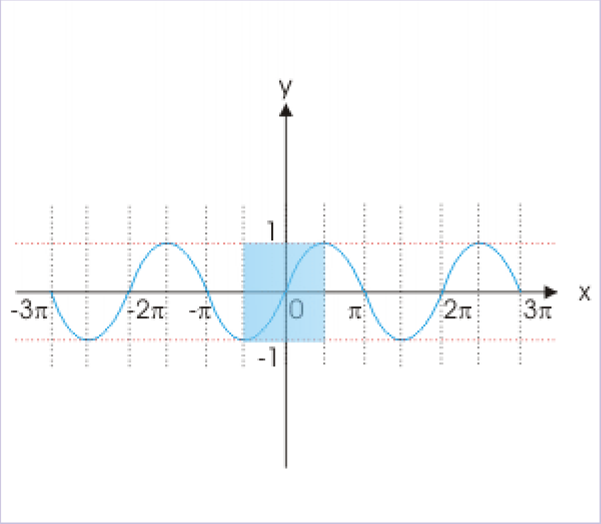
Inverse trigonometric functions

Every angle in the new domain (shortened) is related to a distinct real number in the range. Inversely, every real number in the range is related to a distinct angle in the domain of the trigonometric function. We are aware that the elements of the "ordered pair" in inverse relation exchanges their places. Therefore, it follows that domain and range of trigonometric function are exchanged for corresponding inverse function i.e. domain becomes range and range becomes domain.

arcsine function

The arcsine function is inverse function of trigonometric sine function. From the plot of sine function, it is clear that an interval between $-\pi/2$ and $\pi/2$ includes all possible values of sine function only once. Note that end points are included. The redefinition of domain of trigonometric function, however, does not change the range.

sine function



Redefined domain of function

$$\text{Domain of sine} = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

$$\text{Range of sine} = [-1,1]$$

This redefinition renders sine function invertible. Clearly, the domain and range are exchanged for the inverse function. Hence, domain and range of the inverse function are :

$$\text{Domain of arcsine} = [-1,1]$$

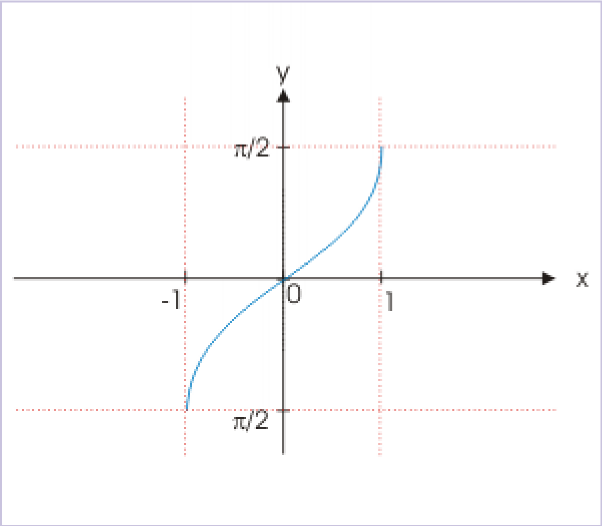
$$\text{Range of arcsine} = \left[-\frac{\pi}{2,}\frac{\pi}{2}\right]$$

Therefore, we define arcsine function as :

$$f : \left[-1,1\right] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \text{by } f(x) = \arcsin(x)$$

The arcsin(x) .vs. x graph is shown here.

arcsine function

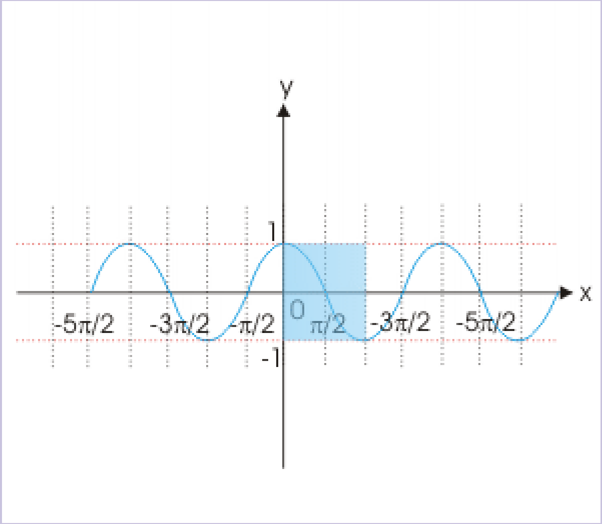


The arcsine function .vs. real value

arccosine function

The arccosine function is inverse function of trigonometric cosine function. From the plot of cosine function, it is clear that an interval between 0 and π includes all possible values of cosine function only once. Note that end points are included. The redefinition of domain of trigonometric function, however, does not change the range.

cosine function



Redefined domain of function

$$\text{Domain of cosine} = [0,\pi]$$

$$\text{Range of cosine} = [-1,1]$$

This redefinition renders cosine function invertible. Clearly, the domain and range are exchanged for the inverse function. Hence, domain and range of the inverse function are :

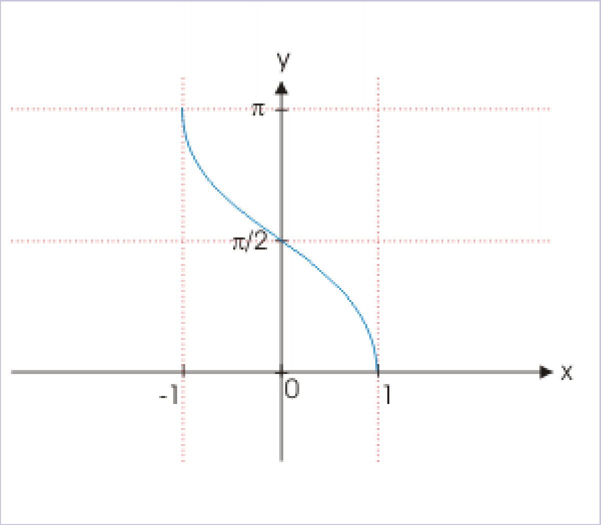
$$\text{Domain of arccosine} = [-1,1]$$

$$\text{Range of arccosine} = [0,\pi]$$

Therefore, we define arccosine function as :

$$f : [-1,1] \rightarrow [0,\pi] \quad \text{by f(x)} = \text{arccos(x)}$$

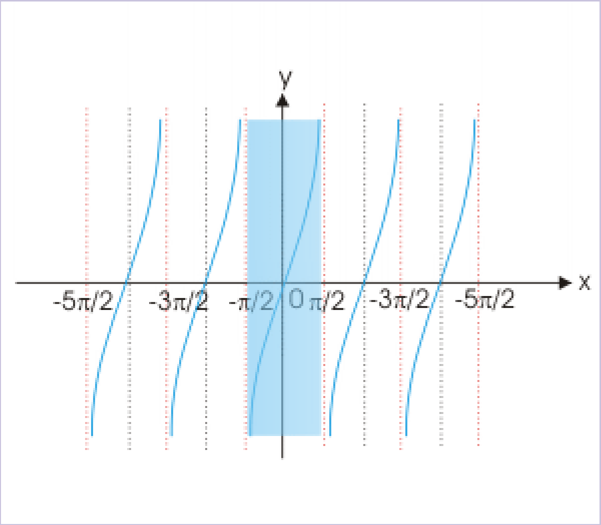
The arccos (x) .vs. x graph is shown here.
arccosine function



The arccosine function .vs. real value

arctangent function

The arctangent function is inverse function of trigonometric tangent function. From the plot of tangent function, it is clear that an interval between $-\pi/2$ and $\pi/2$ includes all possible values of tangent function only once. Note that end points are excluded. The redefinition of domain of trigonometric function, however, does not change the range.
tangent function



Redefined domain of function

$$\text{Domain of tangent} = \left(-\frac{\pi}{2,}\frac{\pi}{2}\right)$$

$$\text{Range of tangent} = R$$

This redefinition renders tangent function invertible. Clearly, the domain and range are exchanged for the inverse function. Hence, domain and range of the inverse function are :

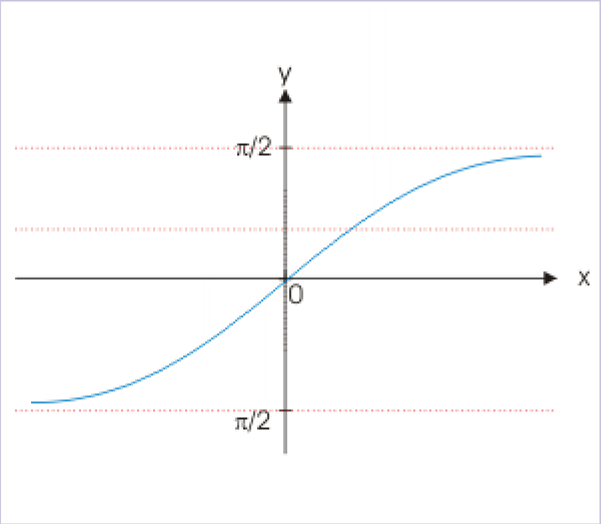
$$\text{Domain of arctangent} = R$$

$$\text{Range of arctangent} = (-\pi/2{,}\pi/2)$$

Therefore, we define arctangent function as :

$$f : R \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad \text{by f}(\text{x}) = \text{arctan}\ (\text{x})$$

The arctan(x) .vs. x graph is shown here.
arctangent function

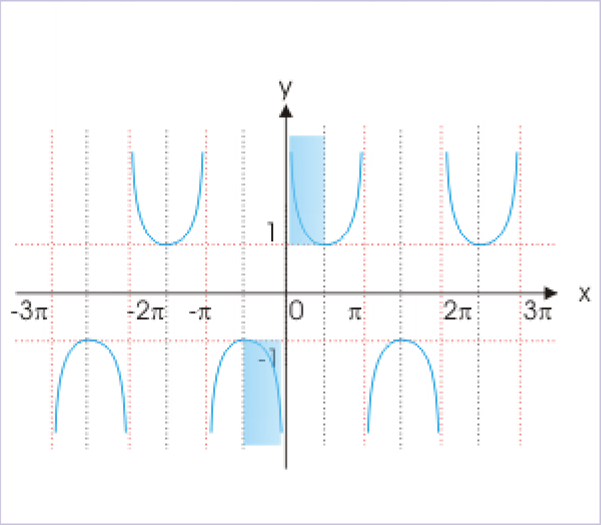


The arctangent function .vs. real value

arccosecant function

The arccosecant function is inverse function of trigonometric cosecant function. From the plot of cosecant function, it is clear that union of two disjointed intervals between “ $-\pi/2$ and 0” and “0 and $\pi/2$ ” includes all possible values of cosecant function only once. Note that zero is excluded, but “ $-\pi/2$ “ and “ $\pi/2$ ” are included . The redefinition of domain of trigonometric function, however, does not change the range.

cosecant function



Redefined domain of function

$$\text{Domain of cosecant} = [-\pi/2, \pi/2] - \{0\}$$

$$\text{Range of cosecant} = (-\infty, -1] \cup [1, \infty) = R - (-1, 1)$$

This redefinition renders cosecant function invertible. Clearly, the domain and range are exchanged for the inverse function. Hence, domain and range of the inverse function are :

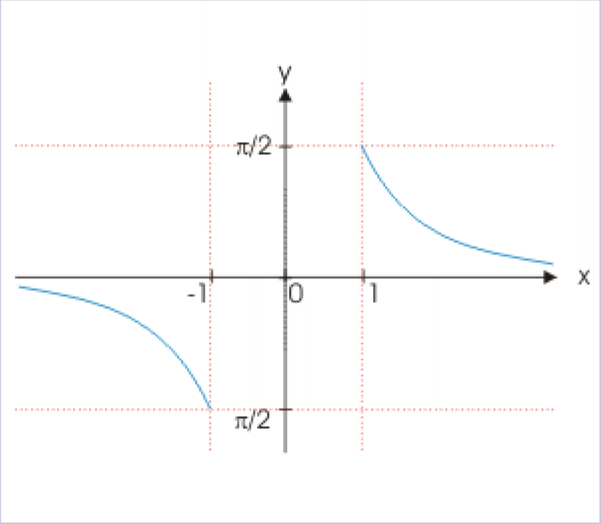
$$\text{Domain of arccosecant} = R - (-1, 1)$$

$$\text{Range of arccosecant} = [-\pi/2, \pi/2] - \{0\}$$

Therefore, we define arccosecant function as :

$$f : R - (-1, 1) \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] - \left\{0\right\} \quad \text{by f(x)} = \text{arccosec (x)}$$

The arccosec(x) .vs. x graph is shown here.
arccosecant function

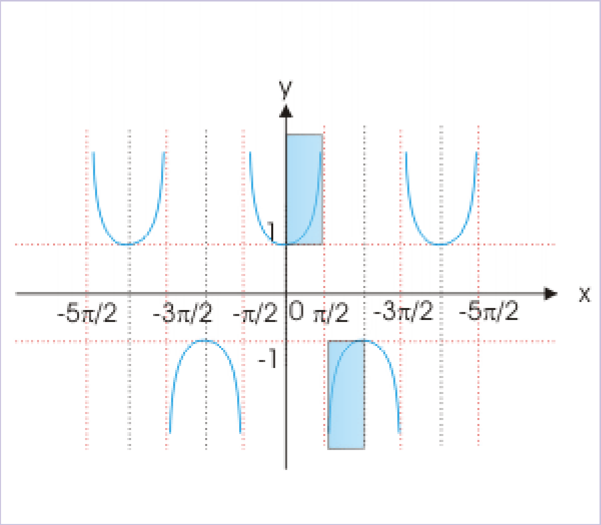


The arccosecant function .vs.
real value

arcsecant function

The arcsecant function is inverse function of trigonometric secant function. From the plot of secant function, it is clear that union of two disjointed intervals between “0 and $\pi/2$ ” and “ $\pi/2$ and π ” includes all possible values of secant function only once. Note that “ $\pi/2$” is excluded. The redefinition of domain of trigonometric function, however, does not change the range.

secant function



Redefined domain of function

$$\text{Domain of secant} = [0,\pi/2) \cup (\pi/2,\pi] = [0,\pi] - \{\pi/2\}$$

$$\text{Range of secant} = (-\infty, -1] \cup [1,\infty) = R - (-1,1)$$

This redefinition renders secant function invertible. Clearly, the domain and range are exchanged for the inverse function. Hence, domain and range of the inverse function are :

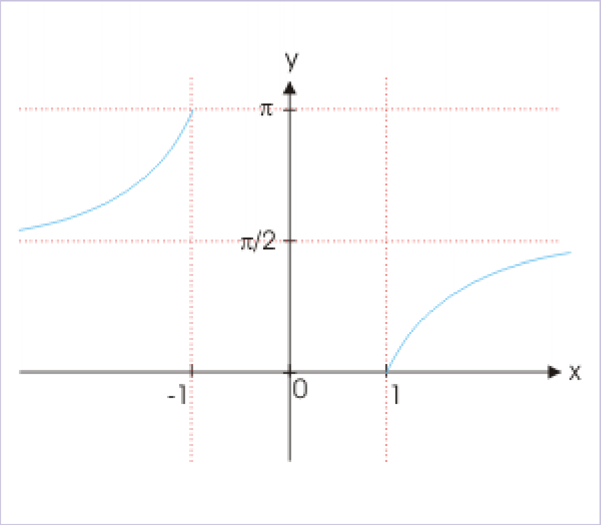
$$\text{Domain of arcsecant} = R - (-1,1)$$

$$\text{Range of arcsecant} = [0,\pi] - \{\pi/2\}$$

Therefore, we define arcsecant function as :

$$f : R - (-1,1) \rightarrow [0,\pi] - \{\pi/2\} \quad \text{by f}(\text{x}) = \text{arcsec}\ (\text{x})$$

The arcsec(x) .vs. x graph is shown here.
arcsecant function

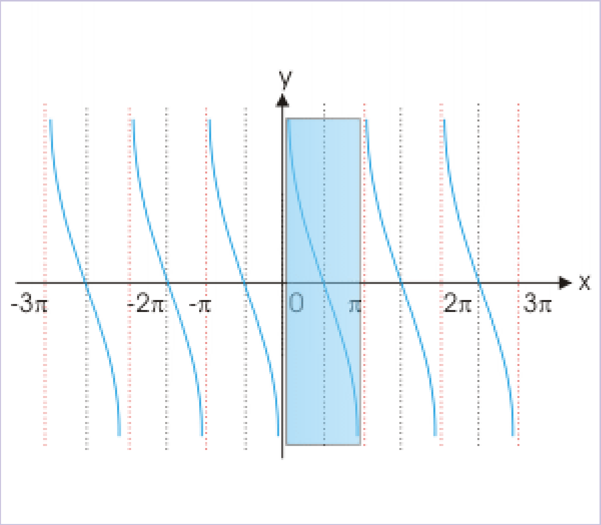


The arcsecant function .vs. real value

arccotangent function

The arccotangent function is inverse function of trigonometric cotangent function. From the plot of cotangent function it is clear that an interval between 0 and π includes all possible values of cotangent function only once. Note that end points are excluded. The redefinition of domain of trigonometric function, however, does not change the range.

cotangent function



Redefined domain of function

$$\text{Domain of cotangent} = (0,\pi)$$

$$\text{Range of cotangent} = R$$

This redefinition renders cotangent function invertible. Clearly, the domain and range are exchanged for the inverse function. Hence, domain and range of the inverse function are :

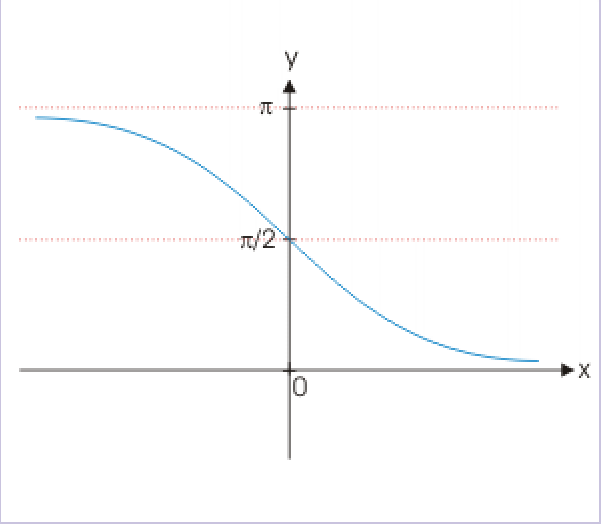
$$\text{Domain of arccotangent} = R$$

$$\text{Range of arccotangent} = (0,\pi)$$

Therefore, we define arccotangent function as :

$$f : R \rightarrow (0,\pi) \quad \text{by f}(\text{x}) = \text{arccot }(\text{x})$$

The arccot(x) .vs. x graph is shown here.
arccotangent function



The arccotangent function .vs.
real value

Example

Example:
Problem : Find y when :

$$y = \tan^{-1}\left(-\frac{1}{\sqrt{3}}\right)$$

Solution : There are multiple angles for which :

$$\Rightarrow \tan y = x = -\frac{1}{\sqrt{3}}$$

However, range of sine function is [-π/2, π/2]. We need to find angle, which falls in this range. Now, acute angle corresponding to the value of 1/√3 is π/6. In accordance with sign diagram, tangent is negative in second and fourth quarters. But range is [-π/2, π/2]. Hence, we need to find angle in fourth quadrant. The angle in the fourth quadrant whose tangent has magnitude of 1/√3 is given by :

$$\Rightarrow y = 2\pi - \frac{\pi}{6} = \frac{11\pi}{6}$$

Corresponding negative angle is :

$$\Rightarrow y = \frac{11\pi}{6} - 2\pi = -\frac{\pi}{6}$$

Example:

Problem : Find domain of the function given by :

$$f(x) = \frac{\cos^{-1}(x)}{[x]}$$

Solution : The given function is quotient of two functions having rational form :

$$f(x) = \frac{g(x)}{h(x)}$$

The domain of quotient is given by :

$$D = D_1 \cap D_2 - \{x : x \quad \text{when h(x)} = 0\}$$

Here, $g(x) = \cos^{-1}(x)$. The domain of arccosine is [-1,1]. Hence,

$$D_1 = \text{Domain of "g"} = [-1,1]$$

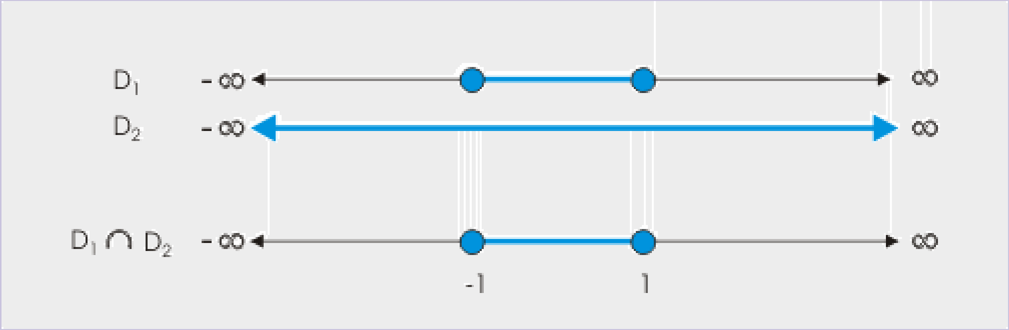
The denominator function h(x) is greatest integer function. Its domain is “R”.

$$D_2 = \text{Domain of "h"} = R$$

The intersection of two domains is :

$$\Rightarrow D_1 \cap D_2 = [-1,1] \cap R = [-1,1]$$

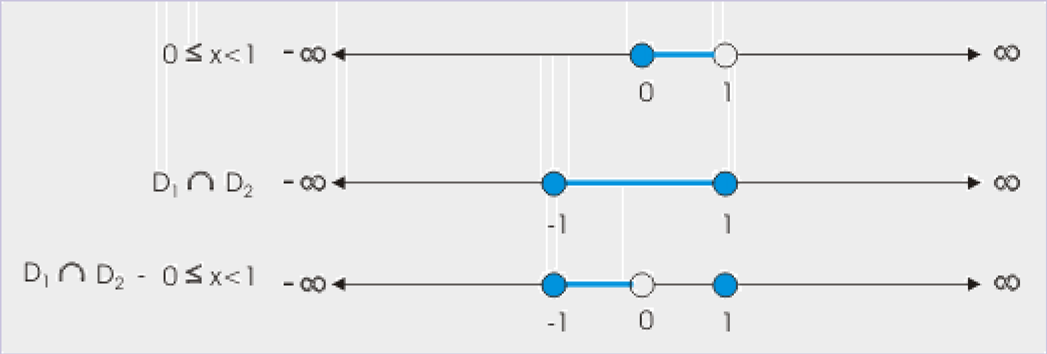
Intersection of domains



The intersection of domains result in common interval.

Now, greatest integer function becomes zero for values of “x” in the interval [0,1). Hence, domain of given function is :

Domain of function



The domain of function is obtained by subtracting interval, which is not permitted.

$$D = D_1 \cap D_2 - [0,1)$$

$$D = [-1,1] - [0,1) = -1 \le x < 0 \cup \{1\}$$

Summary

Redefined domains of trigonometric functions are tabulated here :

```
--------------------------------------------------
-------------------------------------------------
Trigonometric Old New Old New Function Domain
Domain Range Range -------------------------------
--------------------------------------------------
------------------ sine R [-π/2, π/2] [-1,1]
[-1,1] cosine R [0, π] [-1,1] [-1,1] tan R – odd
multiples of π/2 (-π/2, π/2) R R cosecant R –
integer multiple of π [-π/2, π/2] – {0} R – (-1,1)
R – (-1,1) secant R - odd multiples of π/2 [0, π]
– {π/2} R – (-1,1) R – (-1,1) cotangent R –
integer multiple of π (0, π) R R -----------------
--------------------------------------------------
--------------------------------
```

We observe that there is no change in the range – even though domains of the trigonometric functions have changed.

The corresponding domain and range of six inverse trigonometric functions are tabulated here.

```
--------------------------------------------------
------------ Inverse Domain Range Trigonometric
Function -----------------------------------------
--------------------- arcsine [-1,1] [-π/2, π/2]
arccosine [-1,1] [0, π] arctangent R (-π/2, π/2)
arccosecant R – (-1,1) [-π/2, π/2] – {0} arcsecant
R – (-1,1) [0, π] – {π/2} arccotangent R (0, π) --
--------------------------------------------------
----------
```

Exercise

Exercise:

Problem: Find the domain of the function given by :

$$f(x) = 2^{\sin^{-1}(x)}$$

Solution:

The exponent of the exponential function is inverse trigonometric function. Exponential function is real for all real values of exponent. We see here that given function is real for the values of “x” corresponding to which arcsine function is real. Now, domain of arcsine function is [-1,1]. This is the interval of "x" for which arcsine is real. Hence, domain of the given function, “f(x)” is :

$$\text{Domain} = [-1,1]$$

Exercise:

Problem: **Problem :** Find the domain of the function given by :

$$f(x) = \cos^{-1} \frac{3}{3 + \sin x}$$

Solution:

Solution : The given function is an inverse cosine function whose argument is a rational function involving trigonometric function. The domain interval of inverse cosine function is [-1, 1]. Hence, value of argument to inverse cosine function should lie within this interval. It means that :

$$-1 \leq \frac{3}{3 + \sin x} \leq 1$$

Comparing with the form of modulus, $|x| \le 1 \quad \Rightarrow -1 \le x \le 1$, we conclude :

$$\Rightarrow \frac{|3|}{|3+\sin x|} \le 1$$

Since, modulus is a non-negative number, the inequality sign remains same after simplification :

$$\Rightarrow |3| \le |3+\sin x|$$

Again 3 > 0 and 3+sin x > 0, we can open up the expression within the modulus operator without any change in inequality sign :

$$\Rightarrow 3 \le 3+\sin x \quad \Rightarrow 0 \le \sin x \quad \Rightarrow \sin x \ge 0$$

The solution of sine function is the domain of the given function :

$$\text{Domain} = 2n\pi \le x \le (2n+1)\pi, \quad x \in Z$$

Exercise:

Problem: Find range of the function :

$$f(x) = \frac{1}{2-\sin 2x}$$

Solution:

We have already solved this problem by building up interval in earlier module. Here, we shall find domain conventionally by solving for x. The denominator of given function is non-negative as value of sin2x can not exceed 1. Hence, domain of function is real number set R. Further, maximum value of sin2x is 1. Hence,

$$y = f(x) = \frac{1}{2-\sin 2x} > 1$$

This means given function is positive for all real x. Now, solving for x,

$$\Rightarrow 2y - y\sin 2x = 1$$

$$\Rightarrow \sin 2x = \frac{2y-1}{y}$$

$$\Rightarrow x = \frac{1}{2}\sin^{-1}\left(\frac{2y-1}{y}\right)$$

We know that domain of sine inverse function is [-1,1]. Hence,

$$-1 \le \frac{2y-1}{y} \le 1$$

Since y>0, we can simplify this inequality as :

$$-y \le 2y - 1 \le y$$

Either,

$$\Rightarrow 2y - 1 \ge -y$$

$$\Rightarrow y \ge \frac{1}{3}$$

Or,

$$\Rightarrow 2y - 1 \le y$$

$$\Rightarrow y \le 1$$

$$\text{Range} = \left[\frac{1}{3}, 1\right]$$

Exercise:

Problem: Find domain of function

$$f(x) = \sin^{-1}\left\{\log_2\left(x^2 + 3x + 4\right)\right\}$$

Solution:

This is a composite function in which quadratic function is argument of logarithmic function. The logarithmic function is, in turn, argument of inverse sine function. In such case, it is advantageous to evaluate from outer to inner part. The domain of outermost inverse trigonometric function is [-1,1].

$$-1 \le \left\{\log_2\left(x^2 + 3x + 4\right)\right\} \le 1$$

$$\log_2 2^{-1} \le \left\{\log_2\left(x^2 + 3x + 4\right)\right\} \le \log_2 2$$

$$\frac{1}{2} \le \left(x^2 + 3x + 4\right) \le 2$$

For the first inequality,

$$\Rightarrow 2x^2 + 6x + 8 \ge 1$$

$$\Rightarrow 2x^2 + 6x + 7 \ge 0$$

This quadratic function is positive for all value of x. For the second inequality,

$$\Rightarrow x^2 + 3x + 4 \le 2$$

$$\Rightarrow x^2 + 3x + 2 \le 0$$

The solution of this inequality is [1,2]. The intersection of R and [1,2] is [1,2]. Hence, domain of given function is [1,2].

Exercise:

Problem: Find the range of the function

$$f(x) = \cos^{-1}\left(\frac{x^2}{1+x^2}\right)$$

Solution:

Hint : The range of rational expression as argument of inverse trigonometric function is [0,1]. But, domain of arccosine is [-1,1] and range is [0,π]. The function is continuously decreasing. The maximum and minimum values are 0 and1 (see arccosine graph). Hence, range of given function is [0, π/2].

Exercise:

Problem: Find the range of the function

$$f(x) = \operatorname{cosec}^{-1}\left[1+\sin^2 x\right]$$

where [.] denotes greatest integer function.

Solution:

The minimum and maximum value of $\sin^2 x$ is 0 and 1. Hence, range of

$$1+\sin{}^2 x$$

is defined in the interval given by :

$$1 \le 1+\sin^2 x \le 2$$

The corresponding values returned by GIF are 1 and 2. It means :

$$\left[1+\sin^2 x\right] = \{1,2\}$$

But domain of arccosecant is [-π/2, π/2] – {0}. Refer graph of arccosecant. Thus, arccosecant can take only 1 as its argument, which falls within the domain of arccosecant. Hence, range of given function is a singleton :

$$\text{Range} = \left\{\operatorname{cosec}^{-1} 1\right\}$$

Composition of trigonometric function and its inverse

Trigonometric and inverse trigonometric functions are inverse to each other. We can use them to compose new functions. In such composition, trigonometric function represents value of trigonometric ratio, whereas inverse trigonometric function represents angle. The composite function either evaluates to value or angle, depending on particular composition.

Composition representing value of trigonometric function

Sine inverse trigonometric function is given by :

$$y = \sin^{-1} x \quad \Rightarrow x = \sin y \quad \Rightarrow x = \sin \sin^{-1} x$$

$$\Rightarrow \sin \sin^{-1} x = x$$

The composition $\sin \sin^{-1} x$ evaluates to a value. Clearly, x is a value of sine trigonometric function which falls within the range of sine function i.e $x \in \left[-1,1\right]$. It is important to note that domain of inverse function is same as range of the corresponding trigonometric function. We write six compositions denoting value of trigonometric functions as :

$$\sin \sin^{-1} x = x; \quad x \in \left[-1,1\right]$$

$$\cos \cos^{-1} x = x; \quad x \in \left[-1,1\right]$$

$$\tan \tan^{-1} x = x; \quad x \in R$$

$$\cot \cot^{-1} x = x; \quad x \in R$$

$$\sec \sec^{-1} x = x; \quad x \in \left(-\infty, -1\right] \cup \left[1,\infty\right)$$

$$\text{cosec}\, \text{cosec}^{-1} x = x; \quad x \in \left(-\infty, -1\right] \cup \left[1,\infty\right)$$

Composition representing angle

We shall discuss this composition with respect to individual inverse trigonometric ratio.

Composition with arcsine

Sine inverse trigonometric function is given by :

$$y = \sin^{-1} x \quad \Rightarrow x = \sin y \quad \Rightarrow y = \sin^{-1} \sin y$$

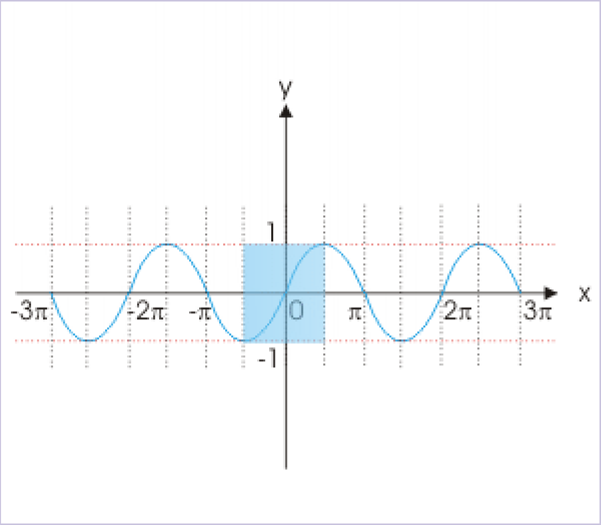
In order to maintain generality, we replace y by x as :

$$\Rightarrow \sin^{-1} \sin x = x$$

The composition $\sin^{-1} \sin x$ evaluates to an angle. Clearly, x is angle value – not the value of trigonometric ratio. However, we know that we use a truncated domain of trigonometric function for defining range of inverse function. The values in the interval are selected such that all unique values of sine trigonometric function are represented. It means that expression on LHS of the equation i.e. $\sin^{-1} \sin x$ evaluates to angle values lying in the interval $[-\pi/2, \pi/2]$.

$$\sin^{-1} \sin x = x; \quad x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

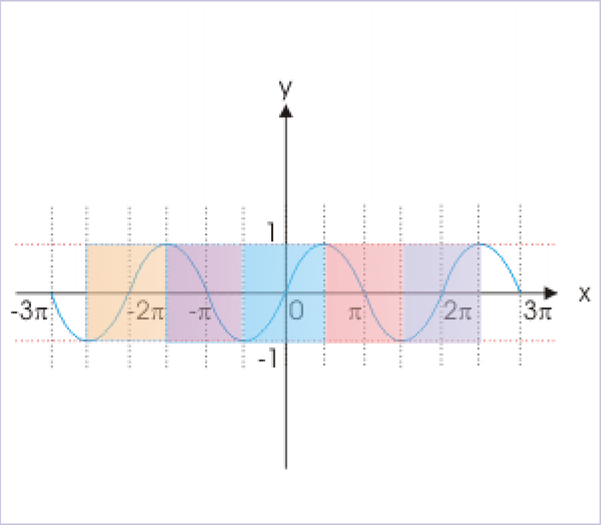
However, x as argument of sine function can assume angle values belonging to real number set. It means angles represented by LHS and RHS can be different if we consider angle values beyond principal set selected to render corresponding trigonometric function invertible.
sine function



Principle domain

Let us consider adjacent intervals such that all sine values are included once. Such intervals are $[\pi/2, 3\pi/2]$, $[3\pi/2, 5\pi/2]$ etc on the right side and $[-3\pi/2, -\pi/2]$, $[-5\pi/2, -3\pi/2]$ etc on the left side of the principal interval.

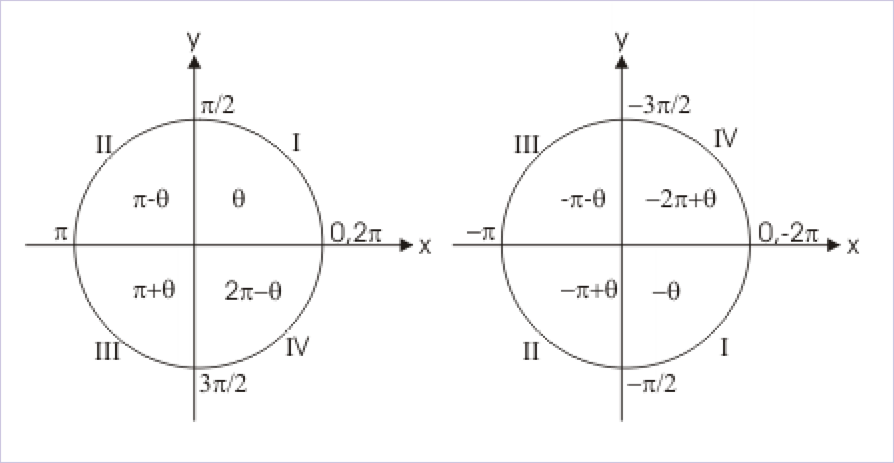
sine function



Additional domains for inversion

Our task now is to determine angles in any of these new intervals, say $[\pi/2, 3\pi/2]$, corresponding to angles in the principal interval. We make use of value diagram which allows to determine angles having same trigonometric values. Let us consider a positive acute angle "θ" in the principal interval. This lies in the first quadrant. The new interval represents second and third quadrants. However, sine is positive in second quadrant and negative in third quadrant. Let the angle corresponding to positive acute angle in principal interval be x. Clearly, x corresponding to positive acute angle θ lies in second quadrant and is given by :

Value diagrams



Value diagrams for positive and negative angles

$$x = \pi - \theta$$

$$\Rightarrow \theta = \pi - x$$

Hence,

$$\Rightarrow \sin^{-1}\sin x = \pi - x; \quad x \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$$

In order to find expression corresponding to negative angle interval $[-3\pi/2, -\pi/2]$, we need to construct negative value diagram. We know that equivalent negative angle is obtained by deducting “-2π” to the positive angle. Thus, corresponding to expression for positive angles in four quadrants, the expression in terms of negative angles are “-θ”,“-π+θ”,“-π-θ” and “-2π+θ” in four quadrants counted in clockwise direction in the value diagram. Now, we estimate from the sine plot that an angle, corresponding to a positive acute angle, θ, in the principal interval, lies in third negative quadrant. Therefore,

$$x = -\pi - \theta$$

$$\Rightarrow \theta = -\pi - x$$

Hence,

$$\Rightarrow \sin^{-1} \sin x = -\pi - x; \quad x \in \left[-\frac{3\pi}{2}, -\frac{\pi}{2}\right]$$

Combining three results,

`|-π-x; x∈ [-3π/2, -π/2] sin⁻¹ sinx = | x; x∈ [-π/2, π/2] | π- x; x∈ [π/2, 3π/2]`

We can similarly find expressions for more such intervals.

Graph of sin^{-1}sinx

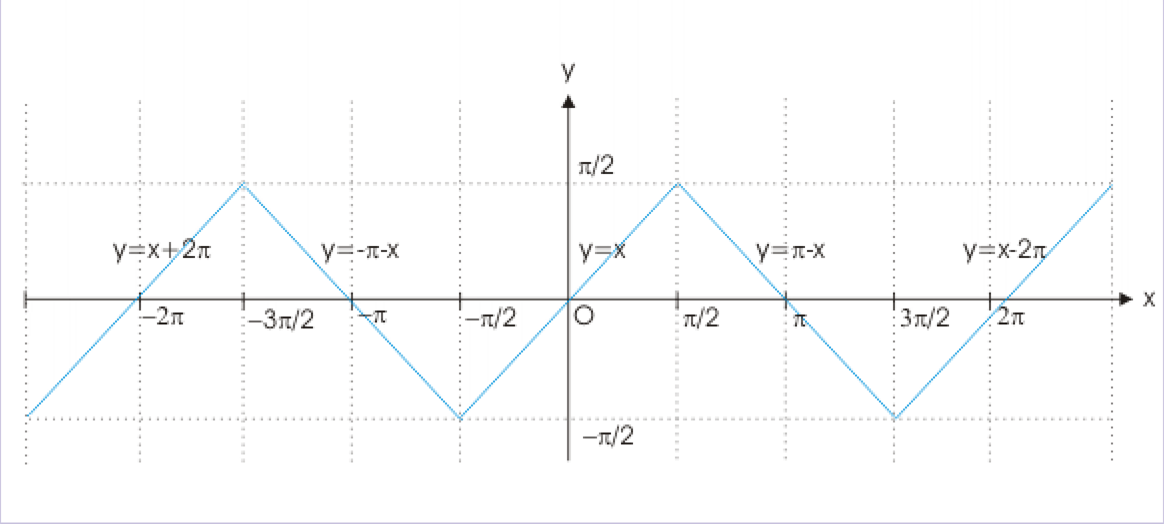
Using three expressions obtained above, we can draw plot of the composition function. We extend the plot, using the fact that composition is a periodic function with a period of 2π. The equation of plot, which is equivalent to plot y=x shifted by 2π towards right, is :

$$y = x - 2\pi$$

The equation of plot, which is equivalent to plot y=x shifted by 2π towards left, is :

$$y = x + 2\pi$$

sine inverse of sine



The function is periodic with period 2π.

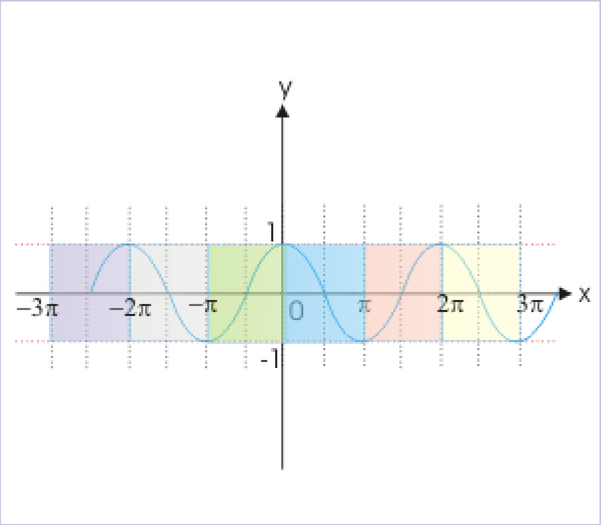
We see that graph of composition is continuous. Its domain is R. Its range is $[-\pi/2, \pi/2]$. The function is periodic with period 2π.

Composition with arccosine

The composition $\cos^{-1} \cos x$ evaluates to angle values lying in the interval [0, π].

$$\cos^{-1} \cos x = x; \quad x \in [0, \pi]$$

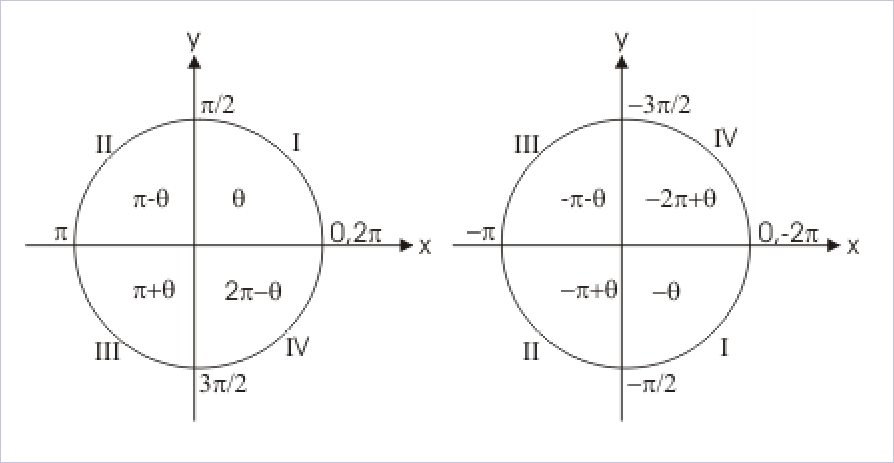
Let us consider adjacent intervals such that all cosine values are included once. Such intervals are [π, 2π], [2π, 3π] etc on the right side and [-π, 0], [-2π, -π] etc on the left side of the principal interval.
cosine function



Additional domains for inversion.

The new interval $[\pi, 2\pi]$ represents third and fourth quadrants. The angle x, corresponding to positive acute angle θ, lies in fourth quadrant. Then,

Value diagrams



Value diagrams for positive and negative angles

$$x = 2\pi - \theta$$

$$\Rightarrow \theta = 2\pi - x$$

Hence,

$$\Rightarrow \cos^{-1}\cos x = 2\pi - x; \quad x \in \left[\pi, 2\pi\right]$$

In order to find expression corresponding to negative angle interval $\left[-\pi, 0\right]$, we estimate from the cosine plot that an angle corresponding to a positive acute angle, θ, in the principal interval lies in first negative quadrant. Therefore,

$$x = -\theta$$

$$\Rightarrow \theta = -x$$

Hence,

$$\Rightarrow \cos^{-1}\cos x = -x; \quad x \in \left[-\pi, 0\right]$$

Combining three results,

`|-x; x∈ [-π, 0] cos⁻¹ cosx = | x; x∈[0, π] |2π- x; x∈ [π, 2π]`

We can similarly find expressions for other intervals.

Graph of cos^{-1}cosx

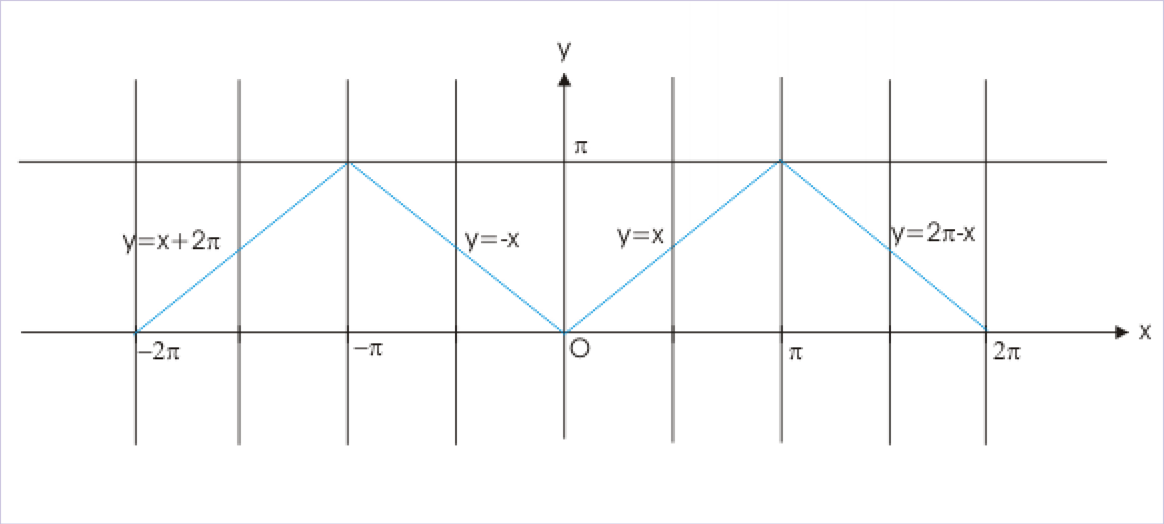
Using three expressions obtained above, we can draw plot of the composition function. We have extended the plot, using the fact that composition is a periodic function with a period of 2π. The equation of plot, which is equivalent to plot y=x shifted by 2π towards right, is :

$$y = x - 2\pi$$

The equation of plot, which is equivalent to plot y=x shifted by 2π towards left, is :

$$y = x + 2\pi$$

cosine inverse of cosine



The function is periodic with period 2π.

We see that graph of composition is continuous. Its domain is R. Its range is $[0,\pi]$. The function is periodic with period 2π.

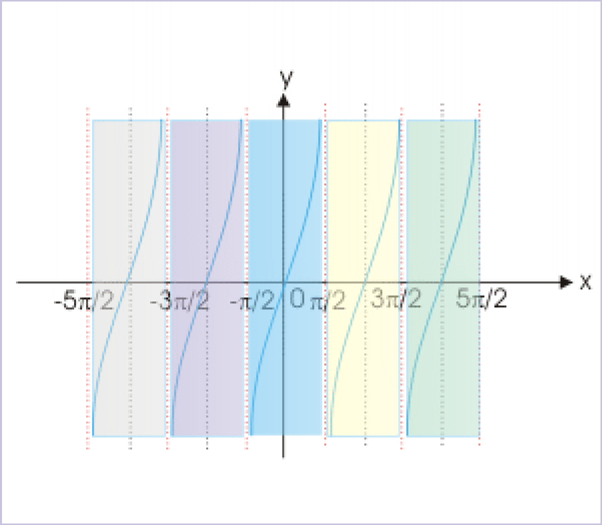
Composition with arctangent

The composition $\tan^{-1}\tan x$ evaluates to angle values lying in the interval $(-\pi/2,\pi/2)$.

$$\tan^{-1}\tan x = x; \quad x \in \left[-\frac{\pi}{2,}\frac{\pi}{2}\right]$$

Let us consider adjacent intervals such that all tangent values are included once. Such intervals are (π/2, 3π/2), (3π/2, 5π/2) etc on the right side and (-3π/2, -π/2), (-5π/2, -3π/2) etc on the left side of the principal interval.

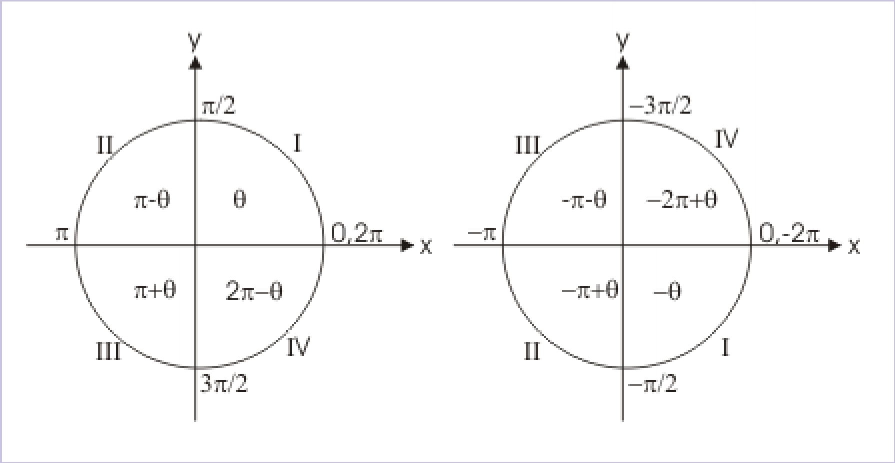
tangent function



Additional domains for inversion.

The new interval $(\pi/2, 3\pi/2)$ represents second and third quadrants. The angle x, corresponding to positive acute angle θ, lies in third quadrant. Then,

Value diagrams



Value diagrams for positive and negative angles

$$x = \pi + \theta$$

$$\Rightarrow \theta = x - \pi$$

Hence,

$$\Rightarrow \tan^{-1}\tan x = x - \pi; \quad x \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$$

In order to find expression corresponding to negative angle interval $(-3\pi/2, -\pi/2)$, we estimate from the tangent plot that an angle corresponding to a positive acute angle, θ, in the principal interval lies in second negative quadrant. Therefore,

$$x = -\pi + \theta$$

$$\Rightarrow \theta = x + \pi$$

Hence,

$$\Rightarrow \tan^{-1}\tan x = x + \pi; \quad x \in \left[-\frac{3\pi}{2,} - \frac{\pi}{2}\right]$$

Combining three results,

`| x+π; x∈ (-3π/2, -π/2) tan⁻¹ tanx = | x; x∈ (-π/2, π/2) | x-π; x∈ (π/2, 3π/2)`

We can similarly find expressions for other intervals.

Graph of tan^{-1}tanx

Using three expressions obtained above, we can draw plot of the composition function. We have extended the plot, using the fact that composition is a periodic function with a period of π. The equation of plot, which is equivalent to plot y=x shifted by π towards right is :

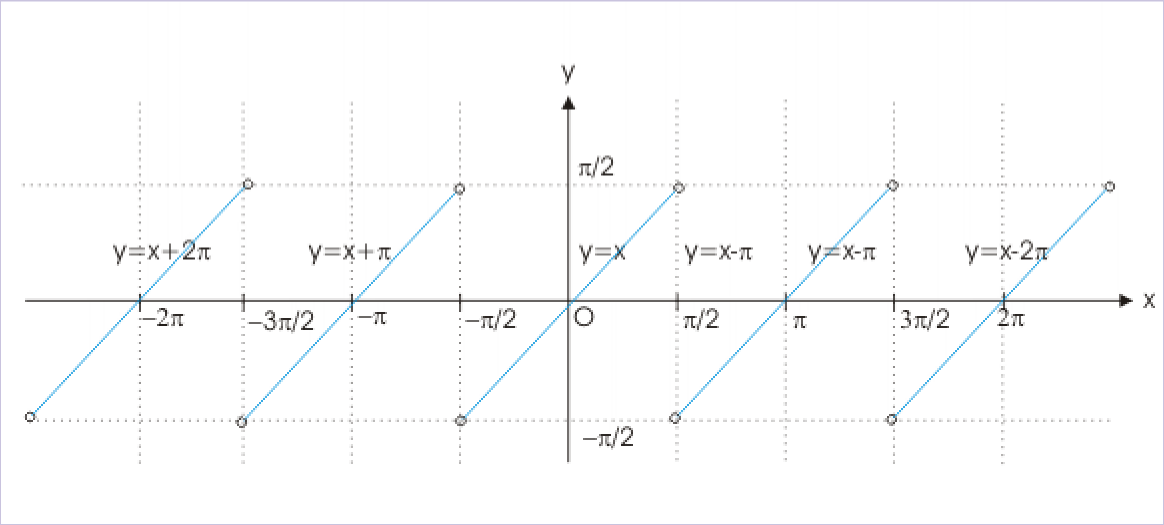
$$y = x - \pi$$

The equation of plot, which is equivalent to plot y=x shifted by π towards left is :

$$y = x + \pi$$

These results are same as obtained earlier. It means that nature of plot is same in the adjacent intervals.

tangent inverse of tangent



The function is periodic with period π.

We see that graph of composition is discontinuous. Its domain is $R - \{(2n+1)\pi/2; n \in Z\}$. Its range is $[-\pi/2,\pi/2]$. The function is periodic with period π.

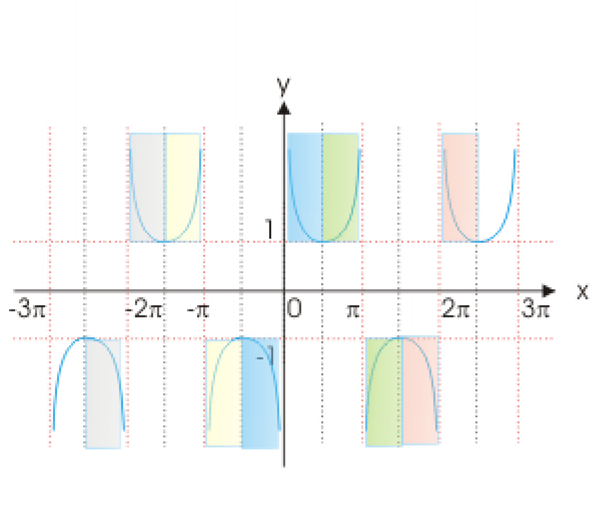
Composition with arccosecant

The composition $\operatorname{cosec}^{-1} \operatorname{cosec} x$ evaluates to angle values lying in the interval $[-\pi/2,\pi/2] - \{0\}$.

$$\operatorname{cosec}^{-1} \operatorname{cosec} x = x; \quad x \in \left[-\frac{\pi}{2,}\frac{\pi}{2}\right] - \left\{0\right\}$$

Let us consider adjacent intervals such that all cosine values are included once. Such intervals are $[\pi/2, 3\pi/2] – \{\pi\}$, $[3\pi/2, 5\pi/2] – \{2\pi\}$ etc on the right side and $[-3\pi/2, -\pi/2] – \{-\pi\}$, $[-5\pi/2, -3\pi/2] – \{-2\pi\}$ etc on the left side of the principal interval.

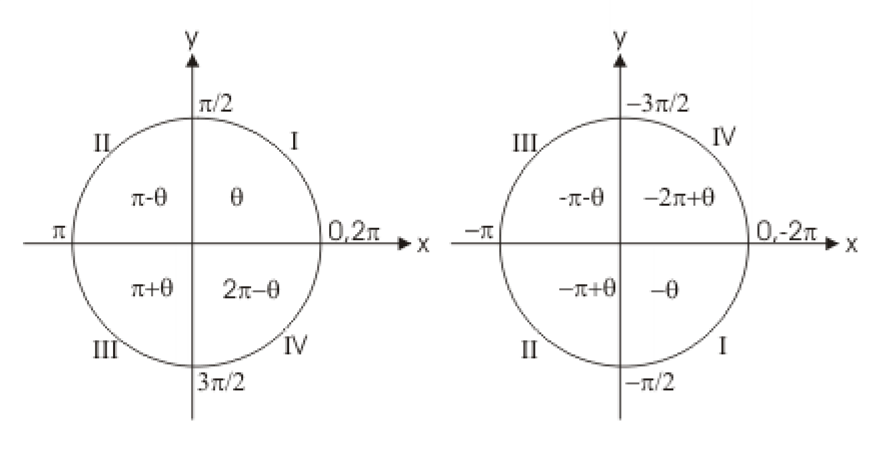
cosecant function



Additional domains for inversion.

The new interval $[\pi/2, 3\pi/2] - \{\pi\}$ lies in second and third quadrants. The angle x corresponding to positive acute angle θ, lies in second quadrant. Then,

Value diagrams



Value diagrams for positive and negative angles

$$x = \pi - \theta$$

$$\Rightarrow \theta = \pi - x$$

Hence,

$$\Rightarrow \operatorname{cosec}^{-1} \operatorname{cosec} x = x; \quad \pi - x \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right] - \left\{\pi\right\}$$

In order to find expression corresponding to negative angle interval $[-3\pi/2, -\pi/2] - \{-\pi\}$, we estimate from the cosecant plot that an angle corresponding to a positive acute angle, θ, in the principal interval lies in third negative quadrant. Therefore,

$$x = -\pi - \theta$$

$$\Rightarrow \theta = -\pi - x$$

Hence,

$$\Rightarrow \operatorname{cosec}^{-1} \operatorname{cosec} x = \pi - x; \quad x \in \left[-\frac{3\pi}{2}, -\frac{\pi}{2}\right] - \left\{-\pi\right\}$$

Combining three results,

`|- π-x; x∈[-3π/2, -π/2] – {-π} cosec⁻¹ cosecx = | x; x∈[- /2, π/2]-{0} | π- x; x∈[π/2, 3π/2] – {π}`

We can similarly find expressions for other intervals.

Graph of cosec^{-1}cosecx

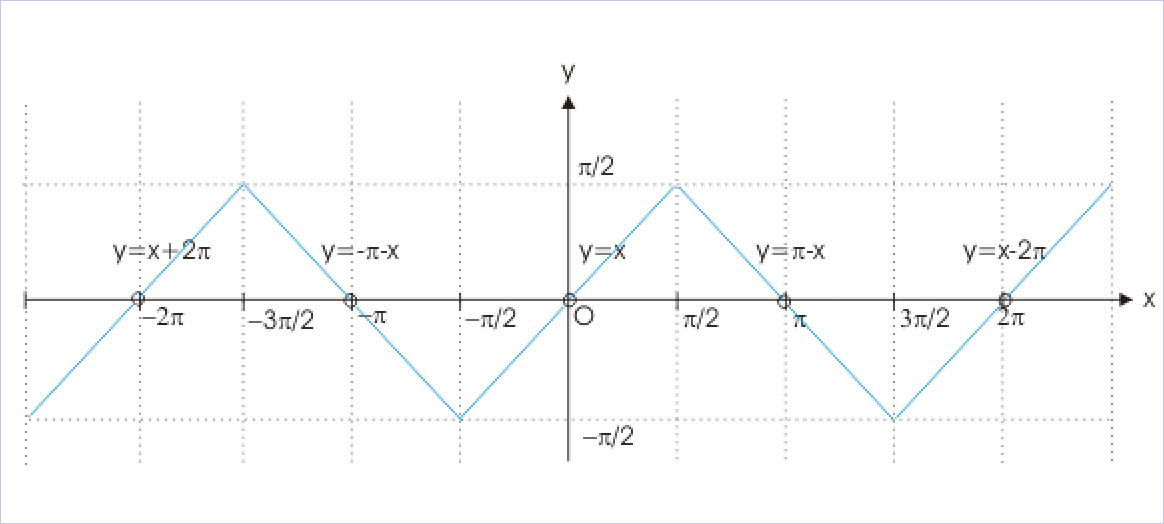
Using three expressions obtained above, we can draw plot of the composition function. We have extended the plot, using the fact that composition is a periodic function with a period of 2π. The equation of plot, which is equivalent to plot y=x shifted by 2π towards right, is :

$$y = x - 2\pi$$

The equation of plot, which is equivalent to plot y=x shifted by 2π towards left, is :

$$y = x + 2\pi$$

cosecant inverse of cosecant



The function is periodic with period 2π.

We see that graph of composition is discontinuous. Its domain is $R - \{n\pi; n \in Z\}$. Its range is $[-\pi/2,\pi/2] - \{0\}$. The function is periodic with period 2π.

Note: We can similarly find out expressions for different intervals for arcsecant and arccotangent compositions. We have left out discussion of these two functions as exercise.

Function operations (exercise)

In this module, we shall work with different function types, which are combined in various ways to form a function. The domain of such functions are determined in accordance with rules for function operations.

Working rules :

- Find domain of the each individual function, which composes the given function. The individual functions may be different function types.
- The domain of a function is unchanged when it is multiplied with a scalar (i.e. a constant)
- The resulting domain after addition, subtraction and multiplication of two functions is given by the intersection of domains, "$D = D_1 \cap D_2$".
- In the case of division, we need to remove values for which denominator is zero.
  $\text{Domain} = D_1 \cap D_2 - \quad \{\text{values of “x” for which denominator is zero}\}$
  .

Note: This exercise module did not follow immediately after the module on function operations. We needed to know different function types first to apply the concept with them.

Problem 1: Find the domain of the function given by :

$$f(x) = \frac{x}{[x-2][x+1]}$$

Solution :

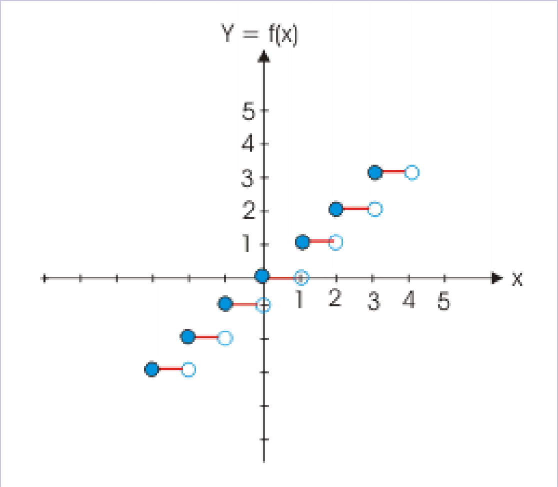
Statement of the problem : The function has rational form. Denominator consists of product of two greatest integer functions.

We can consider, the function as product of three individual functions :

$$f(x) = xX\frac{1}{[x-2]}X\frac{1}{[x+1]}$$

The domain of "x" is “R”. We, now, analyze individual greatest integer functions such that it does not become zero. If we recall the graph of greatest integer function, then we can realize that the value of greatest integer [x] is equal to zero for the interval given by 0≤ x < 1. Following this clue, we find the intervals in which greatest integer functions are zero.

Greatest integer function



The greatest integer function evaluates to zero for 0≤ x < 1.

For $[x-2]=0$,

$$[x-2]=0, \quad \text{if} \quad 0 \le x-2 < 1$$

$$\Rightarrow [x-2]=0, \quad \text{if} \quad 2 \le x < 3$$

$$\Rightarrow [x-2]=0, \quad \text{if} \quad x \in [2,3)$$

It means that given function is undefined for this interval of “x”. The domain of the function for this condition is :

$$D_1 = R - [2,3)$$

Similarly, for $[x+1]=0$

$$[x+1]=0, \quad \text{if} \quad 0 \le x+1 < 1$$

$$\Rightarrow [x+1]=0, \quad \text{if} \quad -1 \le x < 0$$

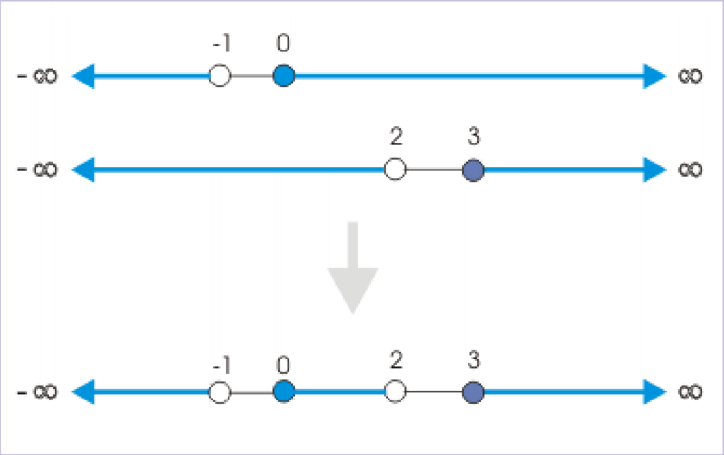
$$\Rightarrow [x+1] = 0, \quad \text{if} \quad x \in [-1,0)$$

The domains for this condition is :

$$D_2 = R - [-1,0)$$

Hence, domain of the given function is intersection of two domains as shown in the figure. Note that we have not considered the domain of numerator, “x”, as its domain is “R” and its intersection with any interval is interval itself.

Domain of the function



The domain of the function is intersection of domains of individual functions.

$$\text{Domain} = D_1 \cap D_2$$

$$\text{Domain} = (-\infty, -1) \quad \cup \quad [0,2) \quad \cup \quad [3,\infty)$$

Problem 2: Find the domain of the function given by :

$$f(x) = \log_e \cos(x-5) + \sqrt{(9-x^2)}$$

Solution :

Statement of the problem : The given function is sum of logarithmic and algebraic function.

Here, we observe that argument (input) of logarithmic function is itself a trigonometric function. We know that cosine function is real for all real values of "x". The important point to realize here is that we have to evaluate logarithmic function for the values of trigonometric function "cos(x-5)" – not for independent variable “x”. Now, the argument of logarithmic function is a positive number. It means that :

$$\cos(x-5) > 0$$

The basic interval of cosine function is $[-\pi/2,\pi/2]$. The solution of the cosine inequality is the domain of the logarithmic function :

$$D_1 = 2n\pi - \frac{\pi}{2} < (x-5) < 2n\pi + \frac{\pi}{2,} \quad n \in Z$$

$$\Rightarrow D_1 = \left(2n - \frac{1}{2}\right)\pi + 5 < x <= \left(2n + \frac{1}{2}\right)\pi + 5, \quad n \in Z$$

For the algebraic function, the expression within the square root is non-negative number :

$$\Rightarrow 9 - x^2 \geq 0 \quad \Rightarrow x^2 - 9 \leq 0 \quad \Rightarrow (x+3)(x-3) \leq 0$$

Clearly, roots of the quadratic equation, when equated to zero, is -3,3. Here, coefficient of quadratic equal equation is positive. Therefore, middle section is negative. Hence, its domain is :

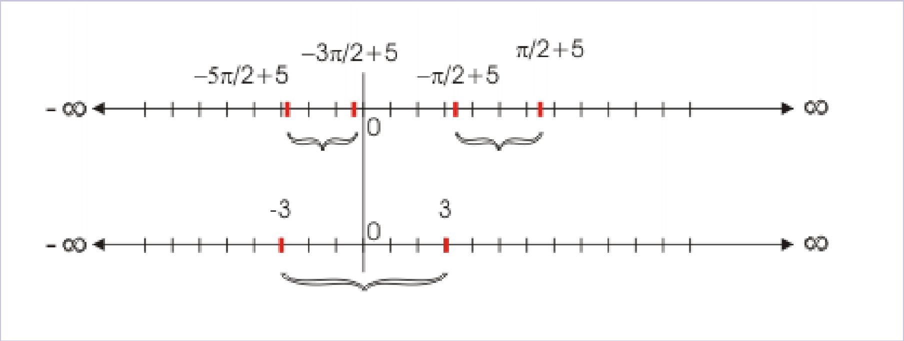
$$D_2 = -3 \leq x \leq 3$$

The domain of given function, f(x), is intersection of two functions i.e.

$$\Rightarrow D = D_1 \cap D_2$$

From the figure, the common interval is between $-5\pi/2 \quad \text{and} \quad -3\pi/2$ as obtained for n = -1.

Domain of the function



The domain of the function is intersection of domains of individual functions.

Note: We should draw a rough number line diagram on paper for few values of “n”.

$$D = \left(-\frac{5\pi}{2}, -\frac{3\pi}{2}\right)$$

Composition of functions (exercise)

Working rules

A. Writing composition :

- Write fog(x) = f(y), where y = g(x)

B. Finding domain of the composition :

- Write down the domain interval of argument function “g(x)” of the composition.
- Write the domain interval of the main function “f(x)” by substituting independent variable by the argument function “g(x)” itself.
- Interpret the interval with argument function, “g(x)”.
- Intersection of two intervals is the valid interval of composition.

> **Note:**This exercise module did not follow immediately after the module on composition of functions. We needed to know different function types first to apply the concept with them.

Problem 1: A function f(x) is given as :

$$f(x) = \{a - x^n\}^{\frac{1}{n}}$$

where a > 0, x > 0 and "n" is a positive integer. Find f{f(x)}

Solution :

Statement of the problem : The domain of the given function is positive number as x>0. In order to find, the composition, we evaluate f(y), where y = f(x).

$$\Rightarrow f\left\{f(x)\right\} = f(y) = (a - y^n)^{\frac{1}{n}} = \left[a - \left\{(a - x^n)^{1/n}\right\}^n\right]^{1/n}$$

$$\Rightarrow f\left\{f(x)\right\} = (a - a - x^n)^{\frac{1}{n}} = (x^n)^{\frac{1}{n}} = x; \quad x > 0$$

Here, composition is that of function with itself. As such, domain of composition is equal to intersection of domain of the given function with itself. But, the intersection of an interval with itself is same interval. Hence, we have retained the domain interval of the composition same as that of given function.

Problem 2: A function f(x) is given as :

$$f(x) = \{2 - x^n\}^{\frac{1}{n}}$$

where x > 0 and "n" is a positive integer. Prove that :

$$f\left\{f(x)\right\} + f\left\{f\left(\frac{1}{x}\right)\right\} \geq 2$$

Solution :

Statement of the problem : The domain of the given function is positive number as x>0. In order to prove the inequality, we need to determine each composition on the left hand side of the given inequality.

We have seen in earlier example that if $f(x) = \{a - x^n\}^{1/n}$, then f{f(x) = x. Hence, if $f(x) = \{2 - x^n\}^{1/n}$, then f{f(x) = x. Similarly, we determine f{f(1/x)}. Here,

$$\Rightarrow f\left\{f\left(\frac{1}{x}\right)\right\} = f(y) = \left[a - \left\{\left(a - \frac{1}{x^n}\right)^{1/n}\right\}^n\right]^{1/n} = \left(a - a - \frac{1}{x^n}\right)^{1/n}$$

$$\Rightarrow f\left\{f\left(\frac{1}{x}\right)\right\} = \left(\frac{1}{x^n}\right)^{\frac{1}{n}} = \frac{1}{x}$$

Substituting these values in the LHS of the inequality, we have :

$$\Rightarrow \text{LHS} = f\left\{f(x)\right\} + f\left\{f\left(\frac{1}{x}\right)\right\} = x + \frac{1}{x}$$

Using algebraic identity $a^2 + b^2 = (a - b)^2 + 2ab$, we have :

$$\Rightarrow \text{LHS} = \left(\sqrt{x} - \frac{1}{\sqrt{x}}\right)^2 + 2$$

But, the square term is a non-negative number. Hence,

$$\Rightarrow \text{LHS} \geq 2$$

$$\Rightarrow f\{f(x)\} + f\{f(1/x)\} \geq 2$$

Problem 3: A function is defined as :

`| -1 ; -2≤x≤0 f(x) = | | x-1 ; 0≤x≤2`

Find composition f(|x|) and its domain.

Solution :

Statement of the problem : The function is defined by different rules in two intervals.

The composition consists of two functions “f(x)” and “|x|”. We know that modulus is defined for all values of “x”. However, domain of “f(x)” is [-2,2]. Hence, domain of composition is intersection of two domains, which is [-2,2]. Here,

`| -1 ; -2≤|x|≤0 and R f(|x|) = | | |x|-1 ; 0≤|x|≤2 and R`

The interval “-2≤[|x|≤0” can be interpreted in parts. The left part is “|x|≥-2”, which is always true. The right part “|x|≤0” is meaningless, which yields no solution for “x”. Therefore, upper interval of the function is not a valid interval. On the other hand, interval “0≤|x|≤2” has two parts. The left part “|x|≥0” is true for all values of “x”. The right part is “|x|≤2”. This expands to :

$$-2 \leq x \leq 2$$

The intersection of “-2≤x≤2” and “R” is “-2≤x≤2”. Hence, composition is :

$$f(|x|) = |x| - 1; \quad -2 \leq x \leq 2$$

Problem 4: Two functions are given as :

$$f(x) = -1 + |x - 1|; \quad -1 \le x \le 3$$

$$g(x) = 2 - |x + 1|; \quad -2 \le x \le 2$$

Find fog(x).

Solution :

Statement of the problem : The domains of the given functions are different. We need to determine composition and interpret domain of the composition.

$$\text{Let} \quad y = g(x) = (2 - |x + 1|); \quad -2 \le x \le 2$$

Now, let us first determine the composition,

$$\Rightarrow fog(x) = f(y) = f(2 - |x + 1|)| = -1 + |2 - |x + 1| - 1|$$

This is the rule of the composition function. In order to find the domain of the composition, we write domains of two functions as an intersection :

$$\Rightarrow fog(x) = -1 + |2 - |x + 1| - 1|; \quad -1 \le y \le 3 \quad \text{and} \quad -2 \le x \le 2$$

We interpret the interval “-1 ≤ y ≤ 3 “ as :

$$-1 \le (2 - |x + 1|) \le 3 \quad \Rightarrow -3 \le (-|x + 1|) \le 1$$

Multiplying each term with “-1” and reversing inequality, we have :

$$3 \ge (|x + 1|) \ge -1$$

But, "|x+1| ≥ -1" is true for all values of “x”. Hence, above inequality is equal to interval given by first part "|x+1| ≤ 3" as :

$$\Rightarrow -3 \le x + 1 \le 3$$

$$\Rightarrow -4 \le x \le 2$$

Hence, interval of the composition is intersection of intervals “-4≤ x ≤2“ and “-2 ≤ x ≤ 2”.

$$\text{Domain} = -2 \le x \le 2$$

The composition, therefore, is :

$$\Rightarrow fog(x) = -1 + |1 - |x+1||; \quad -2 \le x \le 2$$

Problem 5: Two functions are defined as :

$$g(x) = \frac{1-x}{1+x}; \quad 0 \le x \le 1$$

$$h(x) = 4x(1-x); \quad 0 \le x \le 1$$

Determine the composition goh(x) and hog(x).

Solution :

Statement of the problem : The domains of the given functions are same. We need to interpret domain of the composition as we compose the required function.

Let

$$y = h(x) = 4x(1-x) = 4x - 4x^2; \quad 0 <= x <= 1$$

Then,

$$\Rightarrow goh(x) = g(y); \quad 0 \le y \le 1 \quad \text{and} \quad 0 \le x \le 1$$

$$\Rightarrow goh(x) = g\left(4x - 4x^2\right); \quad 0 \le 4x - 4x^2 \le 1 \quad \text{and} \quad 0 \le x \le 1$$

Let us first interpret rule of the function :

$$\Rightarrow goh(x) = \frac{1-y}{1+y} = \frac{1 - \left(4x - 4x^2\right)}{1 + \left(4x - 4x^2\right)} = \frac{1 - 4x + 4x^2}{1 + 4x - 4x^2}$$

Now, we interpret the interval “$0 \le \left(4x - 4x^2\right) \le 1$” in parts. The left part is :

$$\left(4x - 4x^2\right) \ge 0 \quad \Rightarrow x(1-x) \ge 0 \quad \Rightarrow x \ge 0, x \le 1 \quad \Rightarrow 0 \le x \le 1$$

The right part of the interval is :

$$\left(4x - 4x^2\right) \le 1 \quad \Rightarrow \left(4x^2 - 4x\right) \ge -1 \quad \Rightarrow \left(4x^2 - 4x + 1\right) \ge 0$$

Applying sign scheme, this inequality is valid for all values of “x” :

$$\Rightarrow -\infty \le x \le \infty$$

The intersection of two parts is “0≤x≤1”. Thus, interval of composition is intersection of intervals “0≤x≤1” and “0≤x≤1”, which is “0≤x≤1”. Therefore, “goh(x)” is :

$$\Rightarrow goh(x) = \frac{1-4x+4x^2}{1+4x-4x^2}; \quad 0 \le x \le 1$$

For determining hog(x), Let

$$y = g(x) = \frac{1-x}{1+x}; \quad 0 \le x \le 1$$

Then,

$$\Rightarrow hog(x) = g(y); \quad 0 \le y \le 1 \quad \text{and} \quad 0 \le x \le 1$$

$$\Rightarrow hog(x) = g\left(\frac{1-x}{1+x}\right); \quad 0 \le \left(\frac{1-x}{1+x}\right) \le 1 \quad \text{and} \quad 0 \le x \le 1$$

Let us first interpret rule of the function :

$$\Rightarrow hog(x) = 4y(1-y) = 4\frac{1-x}{1+x}\left(1 - \frac{1-x}{1+x}\right)$$

$$\Rightarrow hog(x) = 4\frac{1-x}{1+x}\left(\frac{1+x-1+x}{1+x}\right) = \frac{8x(1-x)}{(1+x)^2}$$

Now, we interpret the interval “$0 \le \left(\frac{1-x}{1+x}\right) \le 1$” in parts. The left part is :

$$\frac{1-x}{1+x} \ge 0 \quad \Rightarrow 1-x \ge 0$$

Applying sign scheme, this inequality is valid for all values of “x” :

$$\Rightarrow 0 \le x \le 1$$

The right part of the interval is :

$$\frac{1-x}{1+x} \le 1 \quad \Rightarrow 1-x \le 1+x \quad \Rightarrow 0 \le 2x \quad \Rightarrow x \ge 0$$

The intersection of two parts is “0≤x≤1”. Thus, interval of composition is intersection of intervals “0≤x≤1” and “0≤x≤1”, which is “0≤x≤1”. Therefore, “goh(x)” is :

$$\Rightarrow hog(x) = \frac{8x(1-x)}{(1+x)^2}; \quad 0 \le x \le 1$$

Problem 6: A function is defined as :

`| 1 + x ; x≥0 f(x) = | | 1- x ; x < 0`

Determine composition f{f(x)}.

Solution :

Statement of the problem : The function is defined by different rules in two intervals.

For f(1+x) in the interval x>=0

`| 1 + 1+ x ; 1+x ≥0 and x≥0 f(1+x) = | | 1- 1 - x ; 1+x < 0 and x≥0`

`| 2+ x ; x ≥-1 and x≥0 f(1+x) = | | - x ; x < -1 and x≥0`

The intersection of upper intervals “x ≥-1 and x≥0” is equal to “x≥0”. There is no common interval for the intersection of lower intervals. Hence,

$$f(1+x) = 2+x; \quad x \ge 0$$

For f(1-x) in the interval x<0

`| 1 + 1- x ; 1-x ≥0 and x≤0 f(1-x) = | | 1- 1 + x ; 1-x < 0 and x<0`

`| 2- x ; x ≤1 and x<0 f(1-x) = | | x ; x > 1 and x<0`

The intersection of upper intervals “x ≤1 and x<0” is equal to “x<0”. There is no common interval for the intersection of lower intervals. Hence,

$$f(1-x) = 2 - x; \quad x < 0$$

Therefore, the composition is :

`| 2+x ; x>0 f{f(x)}= | | 2- x ; x < 0`

Value of a function

The value of a function at “x = a” is denoted as “f(a)”. The working rule for finding value of a function is to replace independent variable “x” by “a”.

Polynomial and rational functions

Problem 1: Find “f(y)”, if

$$y = f(x) = \frac{1-x}{1+x}$$

Solution :

Statement of the problem : The given function is a rational function. We have to evaluate the function when independent variable is function itself.

We need to replace “x” by “y”.

$$\Rightarrow f(y) = \frac{1-y}{1+y} = \frac{1-\frac{1-x}{1+x}}{1+\frac{1-x}{1+x}}$$

$$\Rightarrow f(y) = \frac{1+x-1+x}{1+x+1-x} = \frac{2x}{2} = x$$

Problem 2: Find “f(x)”, if

$$f(x-1) = \left(x^2-1\right)$$

Solution :

Statement of the problem : The given function is a polynomial function with a polynomial as its argument. We have to evaluate the function for independent variable “x”.

We need to replace “x-1” by “x” in the given equation to find “f(x)”. The right hand side expression, however, does not contain term “x-1”. We, therefore, need to find the term, which will replace “x”. Clearly if "x" replaces "x-1", then "x+1" will replace "x-1+1 = x"

Thus, we need to replace “x” by “x+1”.

$$\Rightarrow f(x+1-1) = (x+1)^2 - 1 = x^2 + 2x + 1 - 1 = x^2 + 2x$$

Problem 3: If $\left\{f(x)\right\} = x + \frac{1}{x}$, then prove that :

$$\Rightarrow \left\{f(x)\right\}^3 = f\left(x^3\right) + 3f\left(\frac{1}{x}\right)$$

Solution :

Statement of the problem : The function is a polynomial function. We have to evaluate cube of the function, which involves evaluation of function for arguments, which are independent variable, raised to certain integral powers.

The cube of given function is :

$$\Rightarrow \left\{f(x)\right\}^3 = \left(x + \frac{1}{x}\right)^3 = x^3 + 1/x^3 + 3x^2 X \frac{1}{x} + 3x X \frac{1}{x^2}$$

$$\Rightarrow \left\{f(x)\right\}^3 = x^3 + \frac{1}{x^3} + 3\left(x + \frac{1}{x}\right)$$

Now, $f\left(x^3\right)$ is :

$$f\left(x^3\right) = x^3 + \frac{1}{x^3}$$

Hence,

$$\left\{f(x)\right\}^3 = x^3 + \frac{1}{x^3} + 3\left(x + \frac{1}{x}\right) = f\left(x^3\right) + 3f(x)$$

But, we see that :

$$\Rightarrow f\left(\frac{1}{x}\right) = \frac{1}{x} + x = f(x)$$

Hence,

$$\left\{f(x)\right\}^3 = f\left(x^3\right) + 3f(x) = f\left(x^3\right) + 3f\left(\frac{1}{x}\right)$$

Problem 4: If

$$f(x) = \frac{(1+x)}{(1-x)}$$

Then, find

$$\frac{f(x)f\left(x^2\right)}{1 + \{f(x)\}^2}$$

Solution :

Statement of the problem : The given function is rational function. We have to find the expression which involves (i) function, (ii) function with argument as squared independent variable and (iii) square of the function.

We need to substitute for various terms in the given expression :

$$\Rightarrow \frac{f(x)f\left(x^2\right)}{1 + \{f(x)\}^2} = \frac{\left(\frac{1+x}{1-x}\right) X \left(\frac{1+x^2}{1-x^2}\right)}{1 + \left(\frac{1+x}{1-x}\right)^2}$$

$$= \frac{\frac{(1+x)(1+x^2)}{(1-x)(1-x^2)}}{\frac{(1-x)^2+(1+x)^2}{(1-x)^2}}$$

$$\frac{\frac{(1+x^2)}{(1-x)^2}}{\frac{2(1+x^2)}{(1-x)^2}} = \frac{(1+x^2)}{2(1+x^2)} = \frac{1}{2}$$

Modulus functions

Problem 5: If

$$f(x) = \frac{|x|}{x}; x \neq 0$$

Then, evaluate

$$|f(a) - f(-a)|$$

Solution :

Statement of the problem : Function, f(x), involves modulus and is in rational form. The value of this function, in turn, forms the part of a expression to be evaluated. We have to find the value of expression.

We first evaluate the expression without modulus sign :

$$f(a) - f(-a) = \frac{|a|}{a} - \frac{|-a|}{-a} = \frac{|a|}{a} + \frac{|a|}{a} = \frac{2|a|}{a}; \quad a \neq 0$$

But, we know that $|a| = \pm a$

$$\Rightarrow f(a) - f(-a) = \frac{2X \pm a}{a} = \pm 2$$

Taking modulus of the expression,

$$\Rightarrow |f(a) - f(-a)| = 2; \quad a \neq 0$$

Note that we need to keep the condition for which the given expression is evaluated.

Logarithmic functions

Problem 6: Find $f\left(\frac{2x}{1+x^2}\right)$, if

$$f(x) = \log_e\left(\frac{1+x}{1-x}\right)$$

Solution :

Statement of the problem : The given function is transcendental logarithmic function. We have to evaluate the function for an argument (input to function), which is itself a rational function in independent variable, “x”.

We need to replace “x” by “$2x/1+x^2$”.

$$\Rightarrow f\left(\frac{2x}{1+x^2}\right) = \log_e\left(\frac{1+\frac{2x}{1+x^2}}{1-\frac{2x}{1+x^2}}\right) = \log_e\left(\frac{1+x^2+2x}{1+x^2-2x}\right)$$

$$\Rightarrow f\left(\frac{2x}{1+x^2}\right) = \log_e\left(\frac{1+x}{1-x}\right)^2 = 2\log_e\left(\frac{1+x}{1-x}\right) = 2f(x)$$

Trigonometric functions

Problem 7: Find $f(\pi/4)$, if

$$f(x) = \frac{2\cot x}{1+\cot^2 x}$$

Solution :

Statement of the problem : The given function is a rational function with trigonometric function as independent variable. We have to find the value of function for a particular angle.

We need to replace “x” by “$\pi/4$ ”.

$$\Rightarrow f(\pi/4) = \frac{2\cot(\pi/4)}{1+\cot^2(\pi/4)}$$

As $\cot(\pi/4) = 1$,

$$\Rightarrow f\left(\frac{\pi}{4}\right) = \frac{2X1}{1+1^2} = 1$$

Problem 8: Find $f(\tan\theta)$, if

$$f(x) = \frac{2x}{1+x^2}$$

Solution :

Statement of the problem : The given function is a rational function. We have to evaluate the function for a value, which is itself a trigonometric function.

We need to replace “x” by “$\tan\theta$”.

$$\Rightarrow f(\tan\theta) = \frac{2(\tan\theta)}{1+\tan^2\theta} = \sin 2\theta$$

Problem 9: If $f(x) = \cos\{\log(x)\}$, then prove that :

$$f(xy) + f\left(\frac{x}{y}\right) = 2f(x)f(y)$$

Solution :

Statement of the problem : The given function, f(x) is a trigonometric function, whose input is a logarithmic function. We have to evaluate LHS of the given equation to equate the same to RHS.

Here, we evaluate each term of the left hand side of the equation separately and then combine the result.

$$\Rightarrow f(xy) = \cos\{\log_e(xy)\} = \cos(\log_e x + \log_e y)$$

$$f\left(\frac{x}{y}\right) = \cos\left\{\log_e\left(\frac{x}{y}\right)\right\} = \cos(\log_e x - \log_e y)$$

Substituting in the LHS expression, we have :

$$\Rightarrow \left\{f(xy) + f\left(\frac{x}{y}\right)\right\} = \cos(\log_e x + \log_e y) + \cos(\log_e x - \log_e y)$$

We know that :

$$\cos C + \cos D = 2\cos\left(\frac{C+D}{2}\right)\cos\left(\frac{C-D}{2}\right)$$

Hence,

$$\Rightarrow \left\{f(xy) + f\left(\frac{x}{y}\right)\right\} = 2\cos\left(\frac{\log_e x + \log_e y + \log_e x - \log_e y}{2}\right)\cos\left(\frac{\log_e x + \log_e y - \log_e x + \log_e y}{2}\right)$$

$$\Rightarrow \left\{f(xy) + f\left(\frac{x}{y}\right)\right\} = 2\cos(\log_e x)\cos(\log_e y)$$

$$\left\{f(xy) + f\left(\frac{x}{y}\right)\right\} = 2\cos(\log_e x)\cos(\log_e y) = 2f(x)f(y)$$

Acknowledgment

Author wishes to thank Mr Jay Sicard, Pasadena TX for suggesting valuable typographical correction on the topic.

Transformation of graphs

Transformation of graphs means changing graphs. This generally allows us to draw graphs of more complicated functions from graphs of basic or simpler functions by applying different transformation techniques. It is important to emphasize here that plotting a graph is an extremely powerful technique and method to know properties of a function such as domain, range, periodicity, polarity and other features which involve differentiability of a function. Subsequently, we shall see that plotting enables us to know these properties more elegantly and easily as compared to other analytical methods.

Graphing of a given function involves modifying graph of a core function. We modify core function and its graph, applying various mathematical operations on the core function. There are two fundamental ways in which we operate on core function and hence its graph. We can either modify input to the function or modify output of function.

Broad categories of transformation

- Transformation applied by modification to input
- Transformation applied by modification to output
- Transformation applied by modulus function
- Transformation applied by greatest integer function
- Transformation applied by fraction part function
- Transformation applied by least integer function

We shall cover first transformation in this module. Others will be taken up in other modules.

Important concepts

Graph of a function

It is a plot of values of function against independent variable x. The value of function changes in accordance with function rule as x changes. Graph depicts these changes pictorially. In the current context, both core function and modified function graphs are plotted against same independent variable x.

Input to the function

What is input to the function? How do we change input to the function? Values are passed to the function through argument of the function. The argument itself is a function in x i.e. independent variable. The simplest form of argument is "x" like in function f(x). The modified arguments are "2x" in function f(2x) or "2x-1" in function f(2x-1). This changes input to the function. Important to underline is that independent variable x remains what it is, but argument of the function changes due to mathematical operation on independent variable. Thus, we modify argument though mathematical operation on independent variable x. Basic possibilities of modifying argument i.e. input by using arithmetic operations on x are addition, subtraction, multiplication, division and negation. In notation, we write modification to the input of the function as :

$$\text{Argument/input} = bx + c; \quad b, c \in R$$

These changes are called internal or pre-composition modifications.

Output of the function

A modification in input to the graph is reflected in the values of the function. This is one way of modifying output and hence corresponding graph. Yet another approach of changing output is by applying arithmetic operations on the function itself. We shall represent such arithmetic operations on the function as :

$$af(x) + d; \quad a, d \in R$$

These changes are called external or post-composition modifications.

Arithmetic operations

Addition/subtraction operations

Addition and subtraction to independent variable x is represented as :

$$x+c; \quad c \in R$$

The notation represents addition operation when c is positive and subtraction when c is negative. In particular, we should underline that notation “bx+c” does not represent addition to independent variable. Rather it represents addition/ subtraction to “bx”. We shall develop proper algorithm to handle such operations subsequently. Similarly, addition and subtraction operation on function is represented as :

$$\mathrm{f(x)}+d; \quad d \in R$$

Again, “af(x) + d” is addition/ subtraction to “af(x)” not to “f(x)”.

Product/division operations

Product and division operations are defined with a positive constant for both independent variable and function. It is because negation i.e. multiplication or division with -1 is a separate operation from the point of graphical effect. In the case of product operation, the magnitude of constants (a or b) is greater than 1 such that resulting value is greater than the original value.

$$bx; \quad |b|>1 \quad \text{for independent variable}$$

$$af(x); \quad |a|>1 \quad \text{for function}$$

The division operation is eqivalent to product operation when value of multiplier is less than 1. In this case, magnitude of constants (a or b) is less than 1 such that resulting value is less than the original value.

$$bx; \quad 0<|b|<1 \quad \text{for independent variable}$$

$$af(x); \quad 0 < |a| < 1 \quad \text{for function}$$

Negation

Negation means multiplication or division by -1.

Effect of arithmetic operations

Addition/ subtraction operation on independent variable results in shifting of core graph along x-axis i..e horizontally. Similarly, product/division operations results in scaling (shrinking or stretching) of core graph horizontally. The change in graphs due to negation is reflected as mirroring (across y–axis) horizontally. Clearly, modifications resulting from modification to input modifies core graph horizontally. Another important aspect of these modification is that changes takes place opposite to that of operation on independent variable. For example, when “2” is added to independent variable, then core graph shifts left which is opposite to the direction of increasing x. A multiplication by 2 shrinks the graph horizontally by a factor 2, whereas division by 2 stretches the graph by a factor of 2.

On the other hand, modification in the output of function is reflected in change in graphs along y-axis i.e. vertically. Effects such as shifting, scaling (shrinking or stretching) or mirroring across x-axis takes place in vertical direction. Also, the effect of modification in output is in the direction of modification as against effects due to modifications to input. A multiplication of function by a positive constant greater than 1, for example, stretches the graph in y-direction as expected. These aspects will be clear as we study each of the modifications mentioned here.

Forms of representation

There is a bit of ambiguity about the nature of constants in symbolic representation of transformation. Consider the representation,

$$af(bx+c)+d; \quad a,b,c \in R$$

In this case "a", "b", "c" and "d" can be either positive or negative depending on the particular transformation. A positive "d" means that graph is shifted up. On the other hand, we can specify constants to be positive in the following representation :

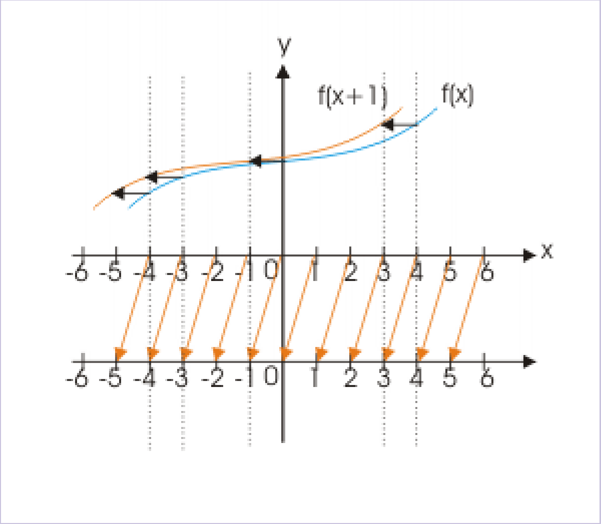
$$\pm af(\pm bx \pm c) \pm d; \quad a,b,c > 0$$

The form of representation appears to be cumbersome, but is more explicit in its intent. It delinks sign from the magnitude of constants. In this case, the signs preceding positive constants need to be interpreted for the nature of transformation. For example, a negative sign before c denotes right horizontal shift. It is, however, clear that both representations are essentially equivalent and their use depends on personal choice or context. This difference does not matter so long we understand the process of graphing.

Transformation of graph by input

Addition and subtraction to independent variable

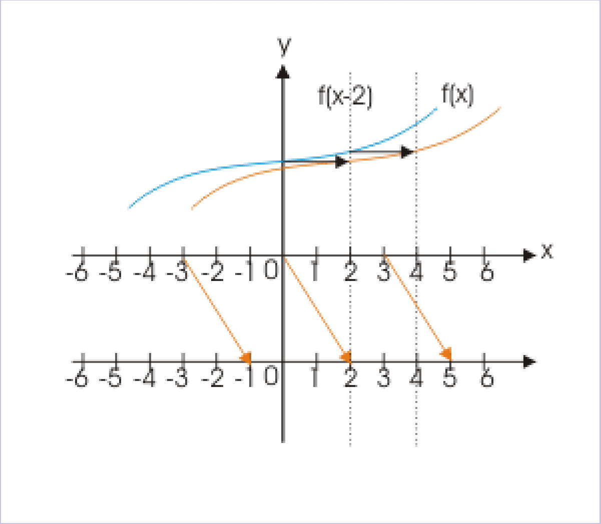
In order to understand this type of transformation, we need to explore how output of the function changes as input to the function changes. Let us consider an example of functions f(x) and f(x+1). The integral values of inedependent variable are same as integral values on x-axis of coordinate system. Note that independent variable is plotted along x-axis as real number line. The integral x+1 values to the function f(x+1) - such that input values are same as that of f(x) - are shown on a separate line just below x-axis. The corresponding values are linked with arrow signs. Input to the function f(x+1) which is same as that of f(x) corresponds to x which is 1 unit smaller. It means graph of f(x+1) is same as graph of f(x), which has been shifted by 1 unit towards left. Else, we can say that the origin of plot (also x-axis) has shifted right by 1 unit.
Shifting of graph parallel to x-axis



Each element of graph is shifted left by same value.

Let us now consider an example of functions f(x) and f(x-2). Input to the function f(x-2) which is same as that of f(x) now appears 2 unit later on x-axis. It means graph of f(x-2) is same as graph of f(x), which has been shifted by 2 units towards right. Else, we can say that the origin of plot (also x-axis) has shifted left by 2 units.

Shifting of graph parallel to x-axis



Each element of graph is shifted

right by same value.

The addition/subtraction transformation is depicted symbolically as :

$$y = f(x) \quad \Rightarrow \quad y = f(x \pm |a|); \quad |a| > 0$$

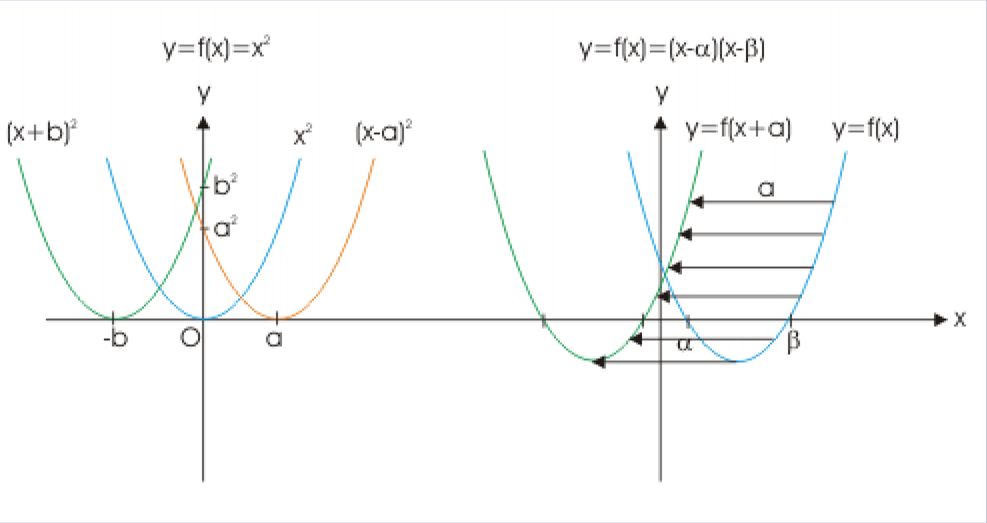
If we add a positive constant to the argument of the function, then value of y at x=x in the new function y=f(x+|a|) is same as that of y=f(x) at x=x-|a|. For this reason, the graph of f(x+|a|) is same as the graph of y=f(x) shifted left by unit “a” in x-direction. Similarly, the graph of f(x-|a|) is same as the graph of y=f(x) shifted right by unit “a” in x-direction.

1 : The plot of y=f(x+|a|); is the plot of y=f(x) shifted left by unit “|a|”.

2 : The plot of y=f(x-|a|); is the plot of y=f(x) shifted right by unit “|a|”.

We use these facts to draw graph of transformed function f(x±a) by shifting graph of f(x) by unit “a” in x-direction. Each point forming the plot is shifted parallel to x-axis (see quadratic graph showm in the of figure below). The graph in the center of left figure depicts monomial function $y = x^2$ with vertex at origin. It is shifted right by “a” units (a>0) and the function representing shifted graph is $y = (x - a)^2$. Note that vertex of parabola is shifted from (0,0) to (a,0). Further, the graph is shifted left by “b” units (b>0) and the function representing shifted graph is $y = (x + b)^2$. In this case, vertex of parabola is shifted from (0,0) to (-b,0).

Shifting of graph parallel to x-axis



Each element of graph is shifted by same value.

Example:

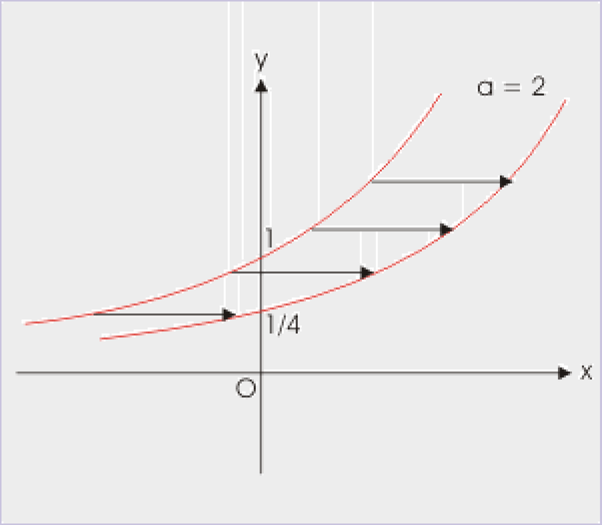
Problem : Draw graph of function $4y = 2^x$.

Solution : Given function is exponential function. On simplification, we have :

$$\Rightarrow y = 2^{-2} X 2^x = 2^{x-2}$$

Here, core graph is $y = 2^x$. We draw its graph first and then shift the graph right by 2 units to get the graph of given function.

Shifting of exponential graph parallel to x-axis



Each element of graph is shifted by same value.

Note that the value of function at x=0 for core and modified functions, respectively, are :

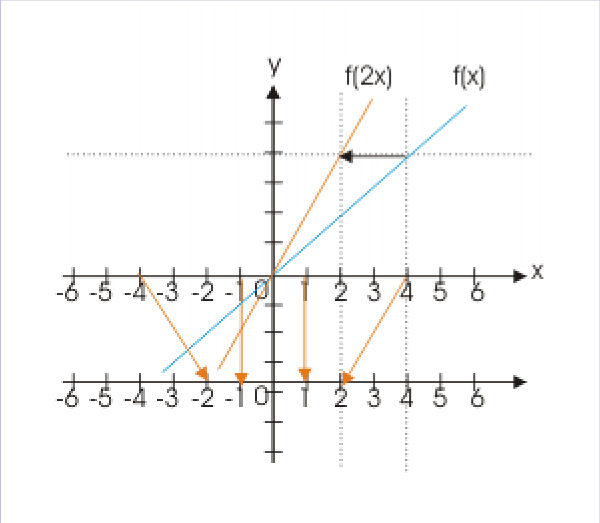
$$\Rightarrow y = 2^x = 2^0 = 1$$

$$\Rightarrow y = 2^{x-2} = 2^{-2} = \frac{1}{4} = 0.25$$

Multiplication and division of independent variable

Let us consider an example of functions f(x) and f(2x). The integral values of independent variable are same as integral values on x-axis of coordinate system. Note that independent variable is plotted along x-axis as real number line. The integral 2x values to the function f(2x) - such that input values are same as that of f(x) - are shown on a separate line just below x-axis. The corresponding values are linked with arrow signs. Input to the function f(2x) which is same as that of f(x) now appears closer to origin by a factor of 2. It means graph of f(2x) is same as graph of f(x), which has been shrunk by a factor 2 towards origin. Else, we can say that x-axis has been stretched by a factor 2.

Multiplication of independent variable

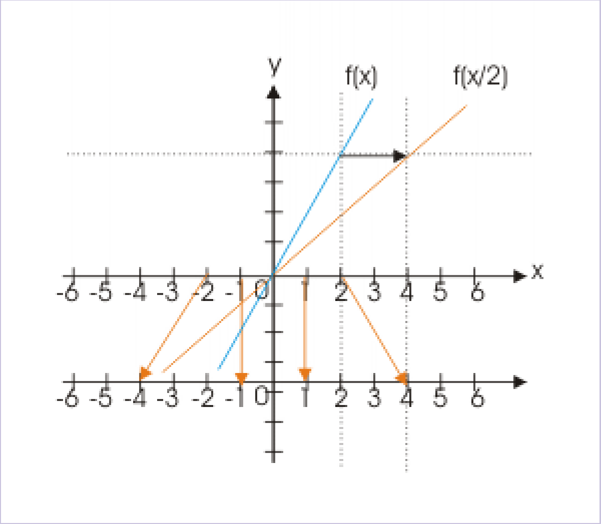


The graph shrinks towards origin.

$$y = f(x) \qquad \Rightarrow y = f(bx); \quad |b| > 1$$

Let us consider another example of functions f(x) and f(x/2). The integral values of independent variable are same as integral values on x-axis of coordinate system. Note that independent variable is plotted along x-axis as real number line. The integral x/2 values to the function f(x/2) - such that input values are same as that of f(x) - are shown on a separate line just below x-axis. The corresponding values are linked with arrow signs. Input to the function f(x/2) which is same as that of f(x) now appears away from origin by a factor of 2. It means graph of f(x/2) is same as graph of f(x), which has been stretched by a factor 2 away from origin. Else, we can say that x-axis has been shrunk by a factor 2.

Multiplication of independent variable



The graph stretches away from origin.

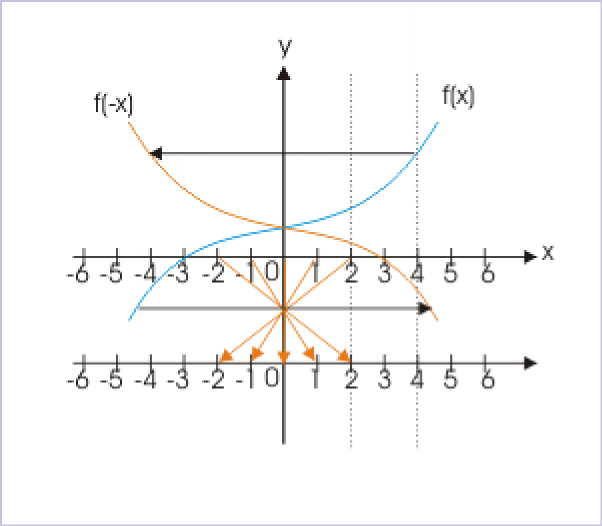
$$y = f(x) \qquad \Rightarrow y = f\left(\frac{x}{b}\right); \quad \left|b\right| > 1$$

Important thing to note about horizontal scaling (shrinking or stretching) is that it takes place with respect to origin of the coordinate system and along x-axis – not about any other point and not along y-axis. What it means that behavior of graph at x=0 remains unchanged. In equivalent term, we can say that y-intercept of graph remains same and is not affected by scaling resulting from multiplication or division of the independent variable.

Negation of independent variable

Let us consider an example of functions f(x) and f(-x). The integral values of independent variable are same as integral values on x-axis of coordinate system. Note that independent variable is plotted along x-axis as real number line. The integral -x values to the function f(-x) - such that input values are same as that of f(x) - are shown on a separate line just below x-axis. The corresponding values are linked with arrow signs. Input to the function f(-x) which is same as that of f(x) now appears to be flipped across y-axis. It means graph of f(-x) is same as graph of f(x), which is mirror image in y-axis i.e. across y-axis.

Negation of independent variable



The graph flipped across y. - axis.

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad y = f(-x)$$

A graph of a function is drawn for values of x in its domain. Depending on the nature of function, we plot function values for both negative and positive values of x. When sign of the independent variable is changed, the function values for negative x become the values of function for positive x and vice-versa. It means that we need to flip the plot across y-axis. In the nutshell, the graph of y=f(-x) can be obtained by taking mirror image of the graph of y=f(x) in y-axis.

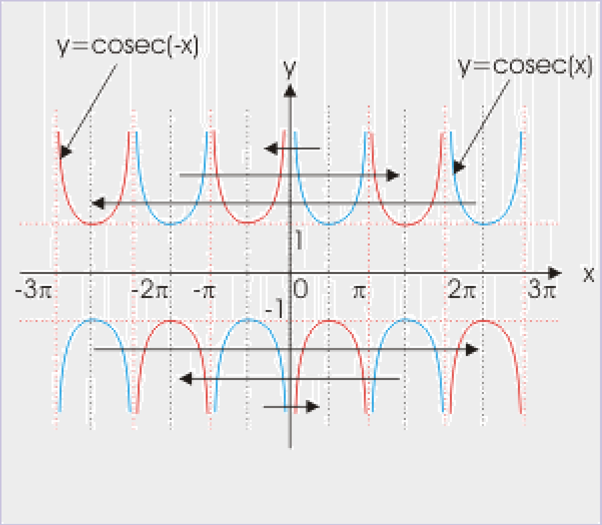
While using this transformation, we should know about even function. For even function. f(x)=f(-x). As such, this transformation will not have any implication for even functions as they are already symmetric about y-axis. It means that two parts of the graph of even function across y-axis are image of each other. For this reason, y=cos(-x) = cos(x), y = |-x|=|x| etc. The graphs of these even functions are not affected by change in sign of independent variable.

Example:

Problem : Draw graph of y=cosec(-x) function

Solution : The plot is obtained by plotting image of core graph y=cosec(x) in y axis.

Changing sign of the argument of graph



The transformed graph is image of core graph in y-axis.

Combined input operations

Certain function are derived from core function as a result of multiple arithmetic operations on independent variable. Consider an example :

$$f(x) = -2x - 2$$

We can consider this as a function composition which is based on identity function f(x) = x as core function. From the composition, it is apparent that order of formation consists of operations as :

(i) f(2x) i.e. multiply independent variable by 2 i.e. shrink the graph horizontally by half.

(ii) f(-2x) i.e. negate independent variable x i.e. flip the graph across y-axis.

(iii) f(-2x-2) i.e. subtract 2 from -2x.

This sequence of operation is not correct for the reason that third operation is a subtraction operation to -2x not to independent variable x, whereas we have defined transformation for subtraction from independent variable. The order of operation for transformation resulting from modifications to input can, therefore, be determined using following considerations :

1 : Order of operations for transformation due to input is opposite to the order of composition.

2 : Precedence of addition/subtraction is higher than that of multiplication/division.

Keeping above two rules in mind, let us rework transformation steps :

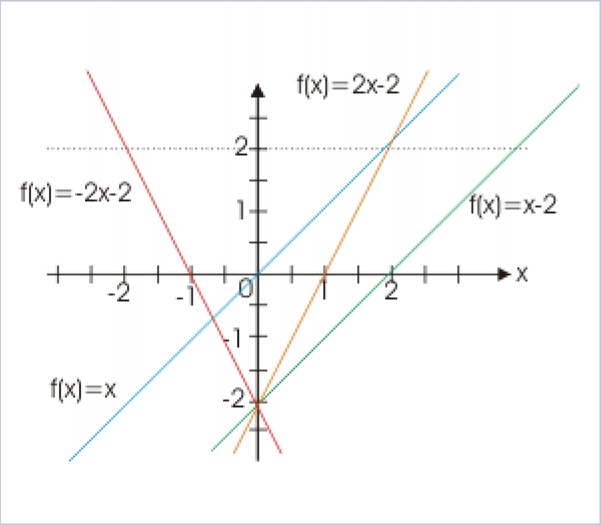
(i) f(x-2) i.e. subtract 2 from independent variable x i.e. shift the graph right by 2 units.

(ii) f(2x-2) i.e. multiply independent variable x by 2 i.e. shrink the graph horizontally by half.

(iii) f(-2x-2) i.e. negate independent variable i.e. flip the graph across y-axis.

This is the correct sequence as all transformations involved are as defined. The resulting graph is shown in the figure below :

Graph of transformed function



Operations are carried in sequence.

It is important the way graph is shrunk horizontally towards origin. Important thing is to ensure that y-intercept is not changed. It can be seen that function before being shrunk is :

$$f(x) = x - 2$$

Its y-intercept is 2. When the graph is shrunk by a factor by 2, the function is :

$$f(x) = 2x - 2$$

The y-intercept is again 2.The graph moves 1 unit half of x-intercept towards origin. Further, we can verify validity of critical points like x and y intercepts to ensure that transformation steps are indeed correct. Here,

$$x = 0, \quad y = -2X0 - 2 = -2$$

$$y = 0, \quad x = -\frac{(y+2)}{2} = -\frac{2}{2} = -1$$

We can decompose a given function in more than one ways so long transformations are valid as defined. Can we rewrite function as y = f{-2(x+1)}? Let us see :

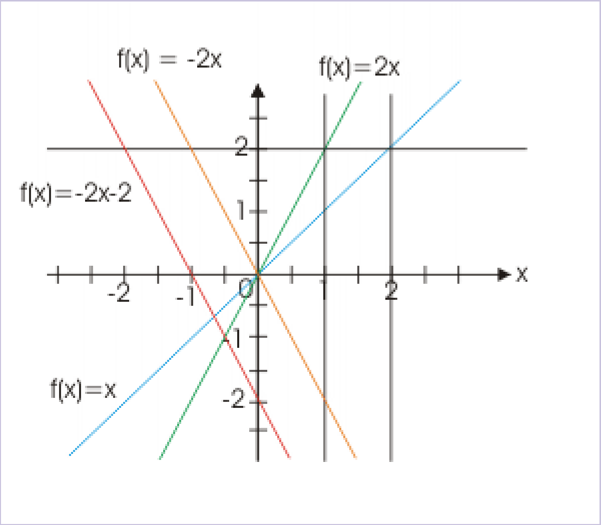
(i) f(2x) i.e. multiply independent variable x by 2 i.e. i.e. shrink the graph horizontally by half.

(ii) f(-2x) i.e. negate independent variable i.e. flip the graph across y-axis.

(iii) f{-2(x+1} i.e. add 1 to independent variable x x i.e. shift the graph left by 1 unit.

This decomposition is valid as transformation steps are consistent with the transformations allowed for arithmetic operations on independent variable.

Graph of transformed function



Operations are carried in sequence.

Horizontal shift

We have discussed transformation resulting in horizontal shift. In the simple case of operation with independent variable alone, the horizontal shift is “c”. In this case, transformation is represented by f(x+c). What is horizontal shift for more general case of transformation represented by f(bx+c)? Let us rearrange argument of the function,

$$f(bx+c) = f\left\{b\left(x+\frac{c}{b}\right)\right\}$$

Comparing with f(x+c), horizontal shift is given by :

$$\text{Horizontal shift} = \frac{c}{b}$$

Transformation of graph by output

We modify output of a function in a couple of ways through arithmetic operations like addition, subtraction, multiplication, division and negation. These operations are similar to the one that we use to modify independent variable. The general symbolic representation for modification to output of a function is represented as :

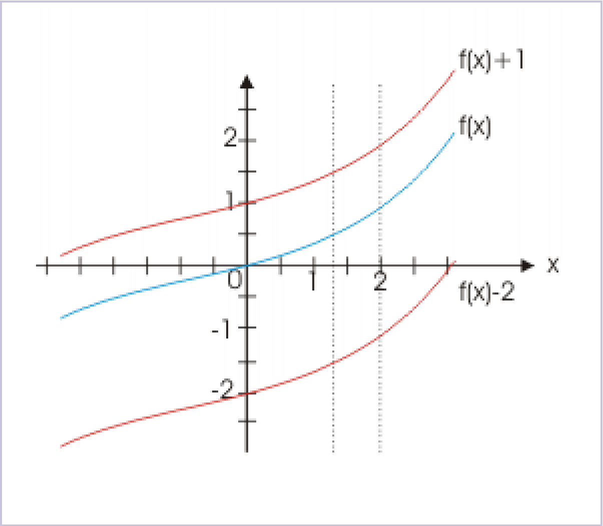
$$af(x) + d; \quad a, d \in R$$

These changes are called external or post-composition modifications. These modifications compliment modifications by input, but in slightly different manner. In the case of modification to output, all effects take place in y-direction i.e. vertical direction as against horizontal transformation arising from modifications affected to input. Second, these transformations are in the direction of operation on output. For example, if we multiply output by a positive constant greater than 1, then graph of core function is stretched along y-axis. This means change in the output is reflected in the same direction in which operation takes place.

Addition and subtraction operation with function

In order to understand this type of transformation, we need to explore how output of the function changes as we add constant value to the output. If we add 1 unit to the function, then each value of function is incremented by 1 unit. It is a straight forward situation. In notation, we would say that the graph of “f(x) + 1” is same as the graph of f(x), which has been moved up by 1 unit. Alternatively, we can also describe this transformation by saying that vertical reference of measurement i.e. x-axis has moved down by 1 unit.

Shifting of graph parallel to y-axis



Each element of graph is shifted by same value.

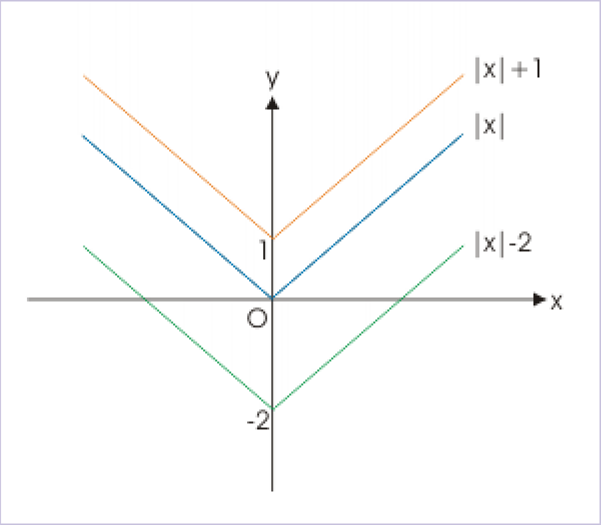
Similarly, if we subtract 1 unit from the function, then each value of function is decremented by 1 unit. In notation, we would say that the graph of “f(x) - 1” is same as the graph of f(x), which has been moved down by 1 unit. Alternatively, we can also describe this transformation by saying that vertical reference of measurement i.e. x-axis has moved up by 1 unit. We conclude :

The plot of y=f(x) + |a|; |a|>0 is the plot of y=f(x) shifted up by unit “a”.

The plot of y=f(x) - |a|; |a|>0 is the plot of y=f(x) shifted down by unit “a”.

We use these facts to draw plot of transformed function f(x±|a|) by shifting plot f(x) by unit “|a|” along y-axis. Each point forming the plot is shifted parallel to x-axis. In the figure below, the plot depicts modulus function y=|x|. It is shifted “1” unit up and the function representing shifted plot is y=|x|+1. Note that corner of plot at x=0 is also shifted by 1 unit along y-axis. Further, the plot is shifted “2” units down and the function representing shifted plot is |x|-2. In this case, corner of plot is shifted by 2 units down along y-axis.

Shifting of graph parallel to y-axis



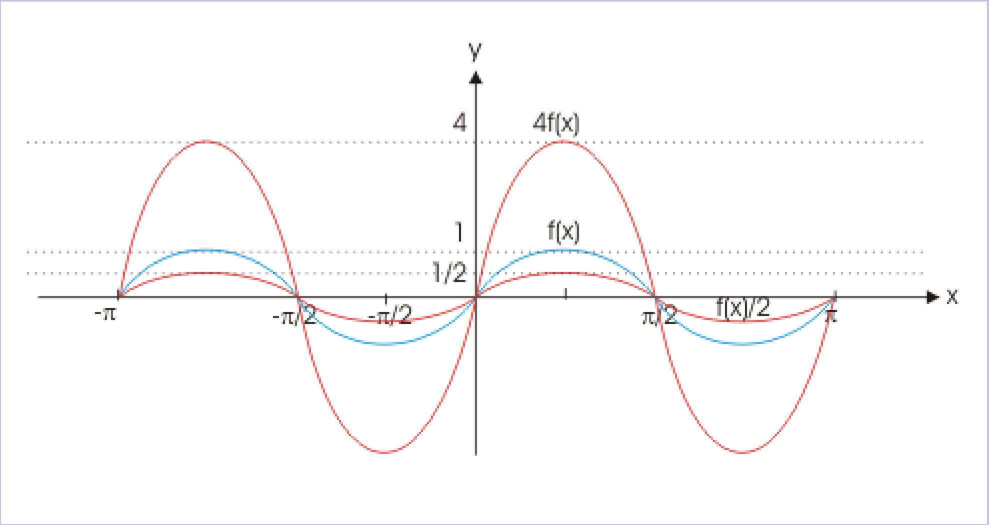
Each element of graph is shifted by same value.

Multiplication and division of function

Multiplication and division scales core graph in accordance with the operation. Scaling, however, is limited to vertical i.e. y-direction. This means modification due to either of these two arithmetic operations has no scaling impact in x-direction. If we multiply output of the function by a positive constant greater than 1, then graph of core function is stretched vertically by the factor, which is equal to the constant being multiplied. The magnification of graph i.e. stretching in y-direction is more noticeable in non-linear graphs like sine and cosine graphs, whose values are bounded in the interval [-1,1]. Let us consider function,

$$y = 4\sin x$$

Scaling of graph



Sine graph is stretched and shrunk.

The amplitude of function "4sinx" is 4 times that of core graph "sinx". In the same fashion, a division by a positive constant greater than 1 results in shrinking of core graph by the factor, which is equal to constant being multiplies. Let us consider division of function :

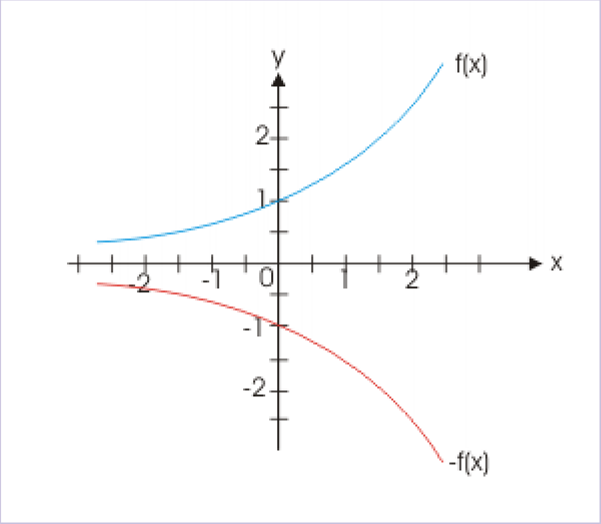
$$y = \frac{1}{2} \sin x$$

The amplitude of the graph "sinx" changes from 1 to 1/2 in the graph of "1/2 sinx".

Negation of function

What would happen if we negate output of a function? Answer is easy. All positive values will turn negative and all negative values will turn positive. It means that graph of core function which is being negated will be swapped across x-axis in the transformation. The graph of “f(-x)”, therefore, is mirror image in x-axis. In other words, we would need to flip the graph f(x) across x-axis to draw graph “-f(x)”.

Changing sign of the graph



The transformed graph is image of the core graph across x-axis.

Example:

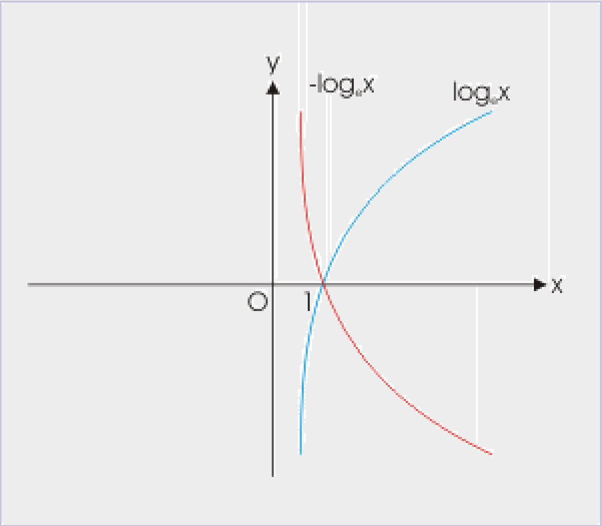
Problem : Draw graph of $y = \log_e \frac{1}{x}$.

Solution : We simplify given function as :

$$\Rightarrow y = \log_e \frac{1}{x} = \log_e 1 - \log_e x = -\log_e x$$

Here, core function is $\mathrm{f}(\mathrm{x}) = \log_e(\mathrm{x})$. Clearly, given function is transformed function of type y=-f(x). We obtain its graph by taking mirror image of the graph of y=f(x) about x-axis. We obtain its graph by taking mirror image of graph of core function about x-axis.

Changing sign of the graph



The transformed graph is image of the core graph about x-axis.

Combined output operations

Certain functions are derived from core function as a result of multiple arithmetic operations on the output of core function. Consider an example :

$$f(x) = -2\sin x - 1$$

We can consider this as a function composition which is based on sine function f(x) = sinx as core function. Here, sequence of operations on the function is important. Difference in interpreting input and output composition is that input composition is evaluated such that defining input transitions are valid. This results in a order of evaluation which gives precedence to addition/subtraction over multiplication/division. This evaluation order is clearly opposite to normal composition order of arithmetic operations in which multiplication/division is given precedence over addition/subtraction. We, therefore, say that decomposition of function for input operation is opposite to that of composition order. In the case of output operation, however, composition order of arithmetic operations is maintained during decomposition. It is logical also. After all, we are operating on a value – not something that goes into function to generate values in accordance with function rule as is the case with independent variable. It is, therefore, expected that we carry out arithmetic operations on the function just the way we evaluate algebraic expressions. In the nutshell, we shall give precedence to multiplication/division over addition/subtraction. In the example abvoe, we subtract "-1" to "-2sinx" - not to core function "sinx".

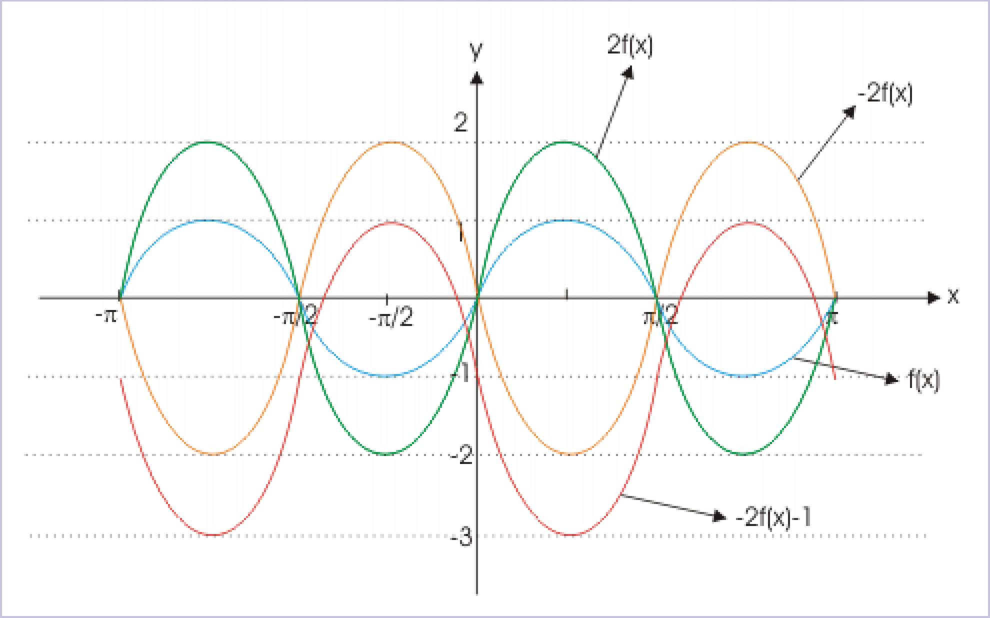
Keeping above in mind, the correct sequence of operation for graphing is :

(i) 2f(x) i.e. multiply function f(x) by 2 i.e. stretch the graph vertically by 2.

(ii) -2f(x) i.e. negate function f(x) i.e. flip the graph across x-axis.

(iii) -2f(x) – 1 i.e. subtract 1 from -2f(x) i.e. shift the graph down by 1 units.

Changing sign of the graph



The transformed graph is image of the base graph about x-axis.

Combined input and output operations

The combined input and output operation is symbolically represented as :

$$af(bx+c)+d; \quad a,b,c,d \in R$$

Carrying out output operation before input operation does not make sense. There will be two different outputs which are not connected to each other. Hence, logical order is that we first carry out input operations then follow it with output operations.

Example:
Problem : Draw $y-1=\log_e(x-2)$
Solution : We rewrite the function :

$$\Rightarrow y=\log_e(x-2)+1$$

In order to plot this function, we plot the graph of core function $y=\log_e x$. Note that when y=0,

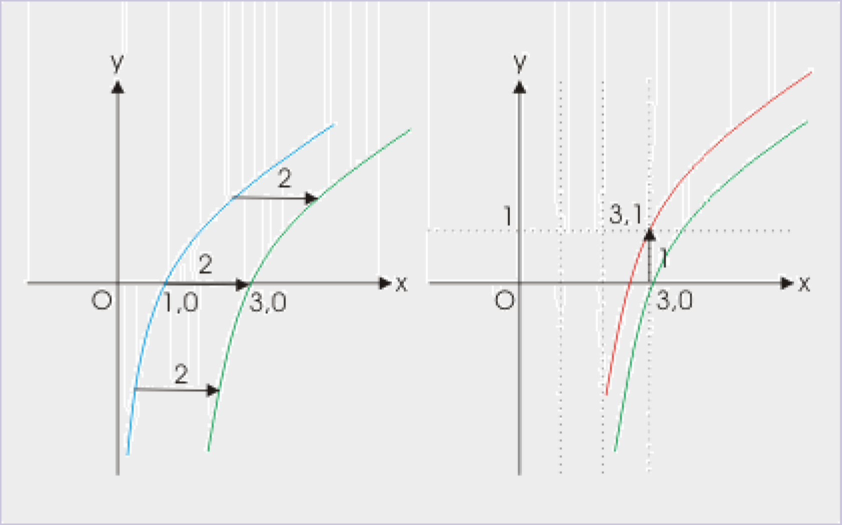
$$y=\log_{e^x}=0 \quad \Rightarrow x=e^0=1$$

In this case, plot intersects x-axis at x=1. Now, the plot of $y=\log_e(x-2)$ is plot of $y=\log_e x$ shifted right by 2 units. Note that when y=0,

$$y=\log e(x-2)=0 \quad \Rightarrow x-2=e^0=1 \quad \Rightarrow x=3$$

The plot of $y=\log_e(x-2)+1$ is plot of $y=\log_e(x-1)$ shifted up by 1 unit.

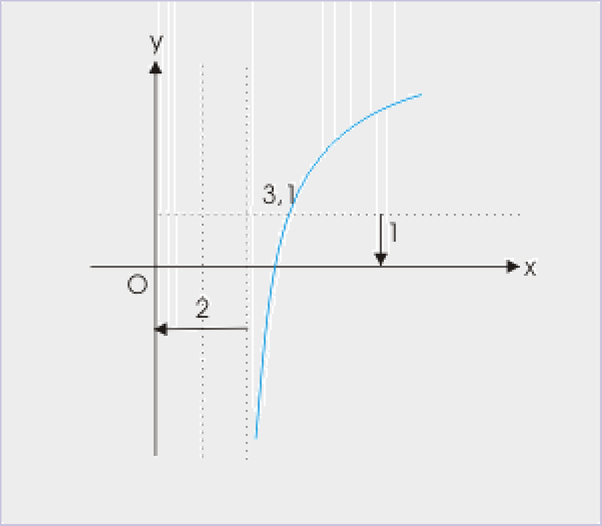
Shifting of logarithmic graph parallel to y-axis



Each element of graph is shifted by same value.

There is yet another alternative to obtain graph of transformed function by shifting axes themselves instead of plot. In the case of shifting either in x or y direction, the operation of shifting graph is equivalent to shifting of axis. Therefore, transformation involving shifting can be affected by shifting axes in opposite directions to that required for the graph. In the example case, we need to move y-axis by 2 units towards left and move x-axis by 1 unit downwards.

Shifting of graph



Each element of graph is shifted
by same value in either
direction.

Example:

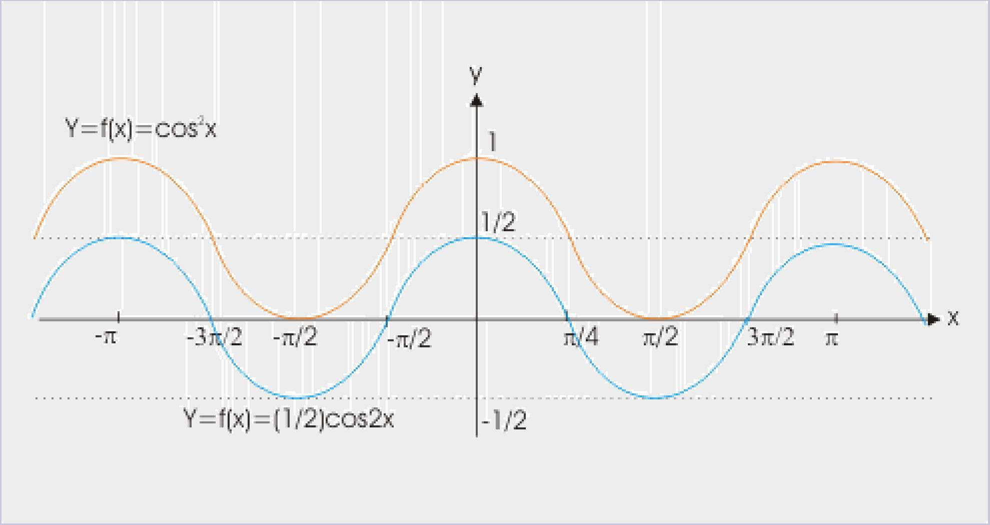
Problem : Draw the plot $y = \cos^2 x$.

Solution : We know that :

$$\Rightarrow y = \cos^2 x = \frac{1 + \cos 2x}{2} = \frac{1}{2} + \frac{\cos 2x}{2}$$

Here, core graph is $y = \cos x$. Multiplying independent variable by 2 shrinks core graph horizontally. As a result its period is reduced from 2π to π as shown in the graph. Division of cos2x by 2 is division operation on function. This operation shrinks the graph cos2x by 2 vertically. Note that amplitude of graph is reduced to 1/2 due to this operation. In the figure, lower graph corresponds to (cos2x)/2. Once we draw graph of (cos2x)/2, we draw given function $y = \cos^2 x$ by shifting the graph of (cos2x)/2 by 1/2 units up.

Graph of squared cosine



Each element of graph is shifted by same value.

Acknowledgment

Author wishes to thank Ms. Aditi Singh, New Delhi for her editorial suggestions.

Transformation of graphs by modulus function

A function like y=f(x) has different elements. We can apply modulus operator to these elements of the function. There are following different possibilities :

1 : $y = f(|x|)$

2 : $y = |f(x)|$

3 : $|y| = f(x)$

4 : $x = |f(y)|$

Plotting concept

The most important point about plotting is to understand that application of modifying operator has different interpretation whether it is applied to independent variable “x” or function definition in x like “f(x)” or it is applied to dependent variable “y” or function definition like “f(y)”. There is a difference in the approach to interpretation.

Clearly, modulus operations have different implications for the graph of f(x). In general, every function can be interpreted to be an operator which operates on its argument, which in itself can be variable like “x”, expression like “$x^2 + 2$ ” or other functions. This role is more visible for functions like modulus, greatest integer, fraction part and least integer function. For this reason, these functions are represented by symbolic notations like | |, [], {} and () as operators.

When operator is applied to independent variable or function definition, we evaluate operation of the operator on independent variable or function value. Here, interpretation is based on “evaluation” of the expression (independent variable or function definition) and application of operator thereafter. This applies to the transformations enumerated at (i) and (ii) above. Consider for example,

$$y = |f(x)|$$

The function of value at any value x=x is first evaluated. Then, modulus of value is calculated. Finally, it is assigned to y as its value.

This basis of interpretation changes when we apply operator to dependent variable “y” or function definition in “y”. Now the basis of interpretation is that of “assigning” a value to a function and then interpreting the assignment. Such is the case with transformations enumerated at (iii) and (iv) above. Consider for example,

$$|y| = f(x)$$

In this case, value of function evaluated at x=x is assigned to modulus function. We interpret equality of the modulus function [y] to a value in accordance with modulus definition. In this case, we know that :

$$|y| = a; a > 0 \quad \Rightarrow y = \pm a$$

$$|y| = a; a = 0 \quad \Rightarrow y = 0$$

$$|y| = a; a < 0 \quad \Rightarrow \text{Modulus can not be equated to negative value. No solution}$$

From the point of view of construction of plot, for a single positive value of f(x), say f(x)=4, we have two values of dependent variable i.e. -4 or 4. This needs to be considered while plotting |y|=f(x). In the plot, values of y are plotted against values of x. In this particular instant, there are two points (4,4) and (4,-4) on the graph corresponding to one value of independent variable (4).

Modulus function applied to the independent variable

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow y = f(|x|)$$

It can be seen that modulus operator here modifies independent varaible of the function. In other words, it is like changing input to the function in accordance with nature of modulus function. The input to the function is now either zero or positive number. This has the implication that part of the graph y=f(x) corresponding to negative value of x is not present in the graph of y=f(|x|). Rather, negative value of x is passed as positive value to the function. This means that negative value of independent variable x yields function value which is equal to function value obtained for corresponding positive x whose magnitude is same as that corresponding negative x. It implies that we can obtain function value for negative x by taking image of positive x across y-axis. This is image in y-axis.

From the point of construction of the graph of y=f(|x|), we need to modify the graph of y=f(x) as :

1 : remove left half of the graph

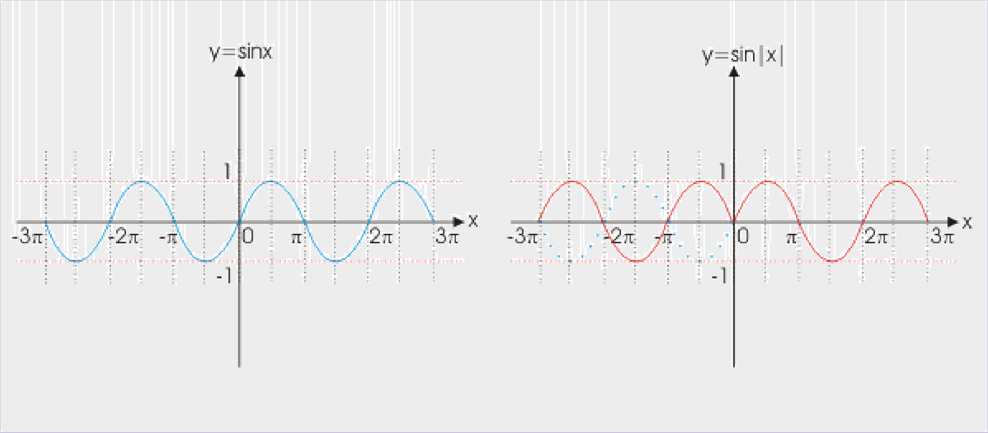
2 : take the mirror image of right half of the graph in y-axis

This completes the construction for y=f(|x|).

Example:
Problem : Draw graph of $y = \sin|x|$.

Solution : First we draw graph of sinx. In order to obtain the graph of y=sin|x|, we remove left half of the graph and take the mirror image of right half of the graph of in y-axis.

Modulus operator applied to sine function



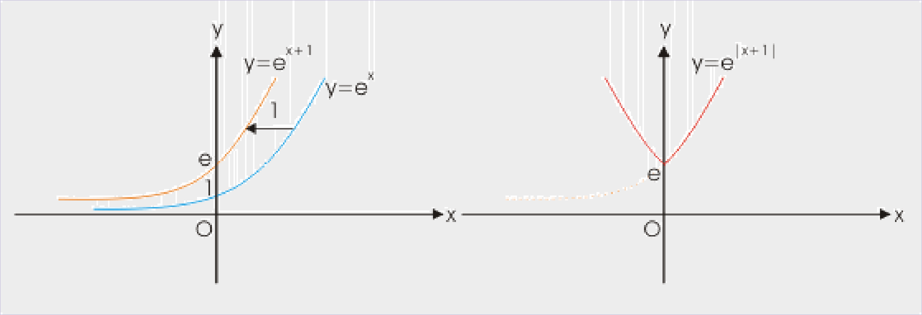
Modulus operator applied to the argument of sine function.

Example:

Problem : Draw graph of $y = e^{|x+1|}$.

Solution : We first draw graph of $y = e^x$. Then, we shift the graph left by 1 unit to obtain the graph of e^{x+1}. At $x = 0$,$y = e^{0+1} = e$. In order to obtain the graph of $y = e^{|x+1|}$, we remove left part of the graph and take the mirror image of right half of the graph of $y = e^{x+1}$ in y-axis.

Modulus operator applied to exponential function



Modulus operator applied to the argument of exponential function.

In order to obtain the graph of $y = e^{|x+1|}$, we remove left part of the graph and take the mirror image of right half of the graph of $y = e^{x+1}$ in y-axis.

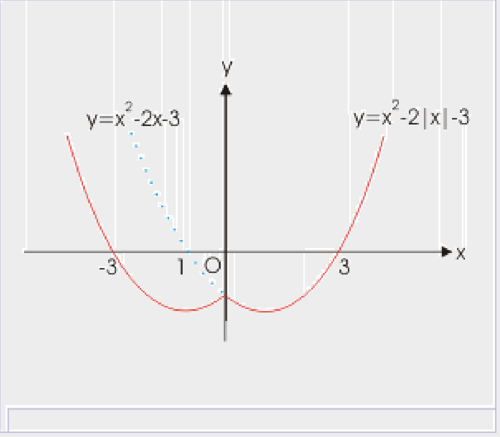
Example:

Problem : Draw graph of $y = x^2 - 2|x| - 3$

Solution : The given expression $f(x) = x^2 - 2|x| - 3$ is obtained by taking modulus of the independent variable of the corresponding quadratic polynomial in x as given here, $f(x) = x^2 - 2x - 3$. Hence, we first draw $f(x) = x^2 - 2x - 3$. The corresponding quadratic equation $f(x) = x^2 - 2x - 3 = 0$ has real roots -1 and 3. The co-efficient of "x^2" is positive. Hence, its plot is a parabola which opens upward and intersects x-axis at x=-1 and x=3.

In order to draw the graph of $f(x) = |x|^2 - 2|x| - 3 = x^2 - 2|x| - 3$, we remove left half of the graph and take the mirror image of right half of the core graph of quadratic function in y-axis.

Modulus operator applied to quadratic function



Modulus operator applied to the quadratic function.

Example:

Problem : Draw graph of function defined as :

$$\Rightarrow y = \frac{1}{|x| + 1}$$

Solution : It is clear that we can obtain given function by applying modulus operator to the independent variable of function given here :

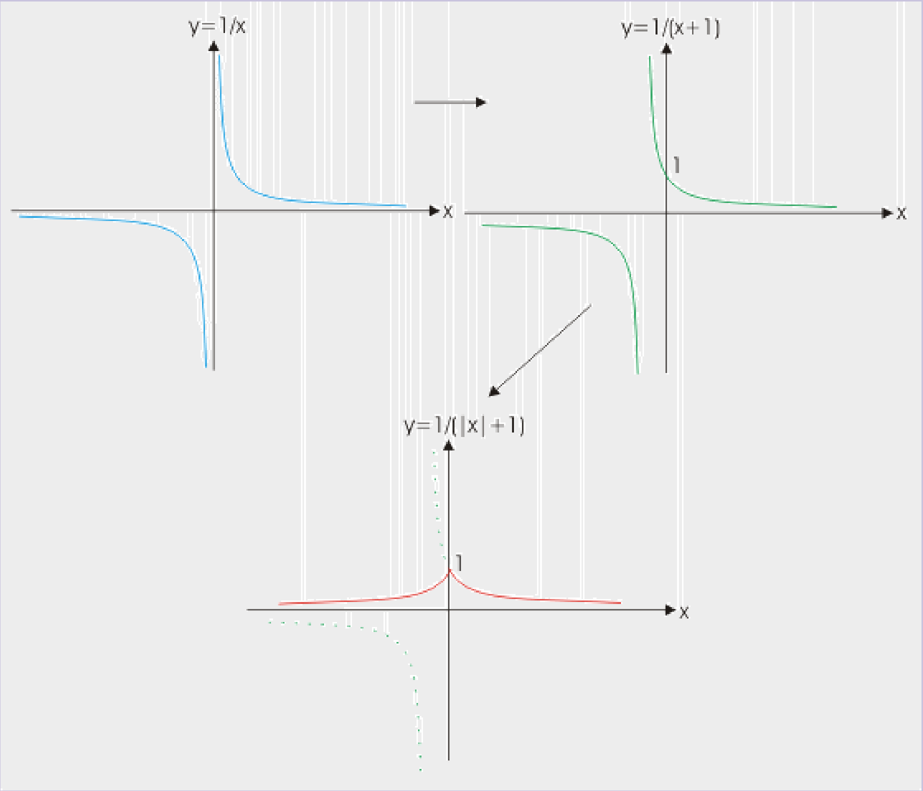
$$\Rightarrow y = \frac{1}{x + 1}$$

This function, in tern, can be obtained by applying shifting modification to the argument of the function given as :

$$\Rightarrow y = \frac{1}{x}$$

We, therefore, first draw $f(x) = 1/x$. Then we draw $g(x) = f(x+1) = 1/(x+1)$ by shifting the graph left by 1 unit. Finally, we draw $h(x) = g(|x|) = 1/(|x|+1)$ by removing left half of the graph and taking mirror image of right half of the graph in y-axis. .

Modulus operator applied to rational function



Modulus operator applied to the argument of rational function.

Modulus function applied to the function

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad y = |f(x)|$$

It can be seen that modulus operator here modifies the value of the function itself. In other words, it is like changing output of the function in accordance with nature of modulus function. The output of the function is now either zero or positive number. This has the implication that part of the graph y=f(x) corresponding to negative function values is not present in the graph of y=|f(x)|. Rather, negative function value of f(x) is converted to positive function value. This change in the sign of function takes place without changing magnitude of the value. It implies that we can obtain function values, which correspond to negative function value in y=f(x) by taking image of negative function values across x-axis. This is image in x-axis.

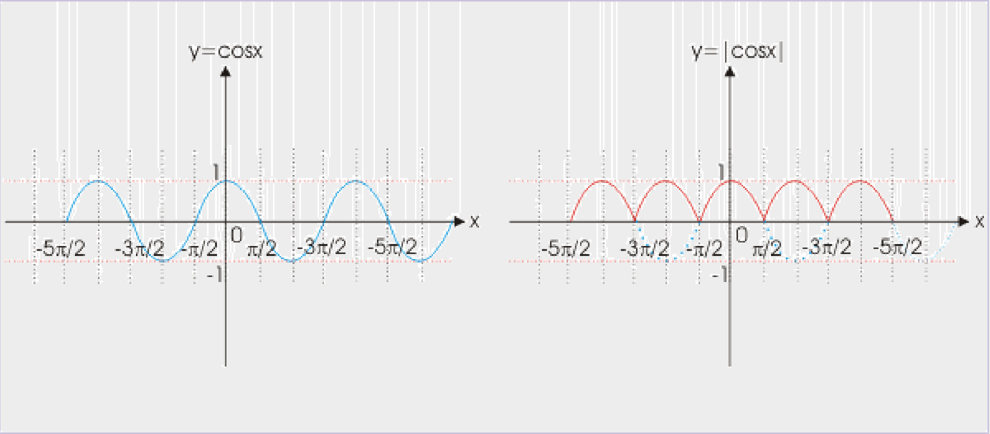
From the point of construction of the graph of y=|f(x)|, we need to modify the graph of y=f(x) as :

(i) take the mirror image of lower half of the graph in x-axis

(ii) remove lower half of the graph

This completes the construction for y=|f(x)|.

Example:
Problem : Draw graph of $y=|\cos x|$.
Solution : We first draw the graph of $y=\cos x$. Then, we take the mirror image of lower half of the graph in x-axis and remove lower half of the graph to complete the construction of graph of $y=|\cos x|$
Modulus operator applied to cosine function

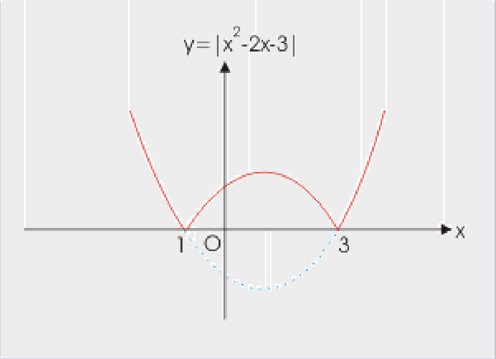


Modulus operator applied to cosine function.

Example:
Problem : Draw graph of $y=|x^2-2x-3|$

Solution : We first draw graph $y = x^2 - 2x - 3$. The roots of corresponding quadratic equation are -1 and 3. After plotting graph of quadratic function, we take the mirror image of lower half of the graph in x-axis and remove lower half of the graph to complete the construction of graph of $y = \left|x^2 - 2x - 3\right|$

Modulus operator applied to quadratic function



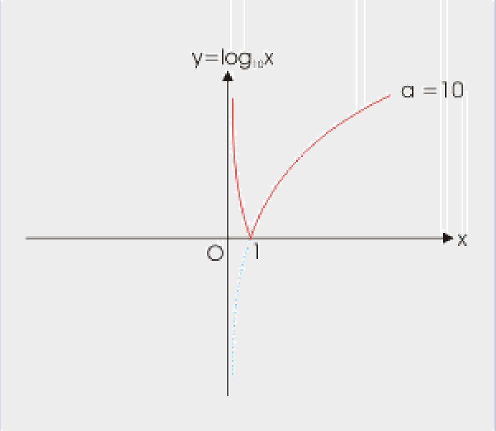
Modulus operator applied to quadratic function.

Example:

Problem : Draw graph of $y = \left|\log_{10} x\right|$.

Solution : We first draw graph $y = \log_{10} x$. Then, we take the mirror image of lower half of the graph in x-axis and remove lower half of the graph to complete the construction of graph of $y = \left|\log_{10} x\right|$.

Modulus operator applied to logarithmic function



Modulus operator applied to logarithmic function.

Modulus function applied to dependent variable

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad |y| = f(x)$$

As discussed in the beginning of module, value of function is first calculated for a given value of x. The value so evaluated is assigned to the modulus function |y|. We interpret assignment to |y| in accordance with the interpretation of equality of the modulus function to a value. In this case, we know that :

$$|y| = f(x); f(x) > 0 \quad \Rightarrow \quad y = \pm f(x)$$

$$|y| = f(x); f(x) = 0 \quad \Rightarrow \quad y = 0$$

$$|y| = f(x); f(x) < 0 \quad \Rightarrow \text{Modulus can not be equated to negative value. No solution}$$

Clearly, we need to neglect all negative values of f(x). For every positive value of f(x), there are two values of dependent expressions -f(x) and f(x). It means that we need to take image of upper part of the graph across x-axis. This is image in x-axis.

From the point of construction of the graph of |y|=f(x), we need to modify the graph of y=f(x) as :

1 : remove lower half of the graph

2 : take the mirror image of upper half of the graph in x-axis

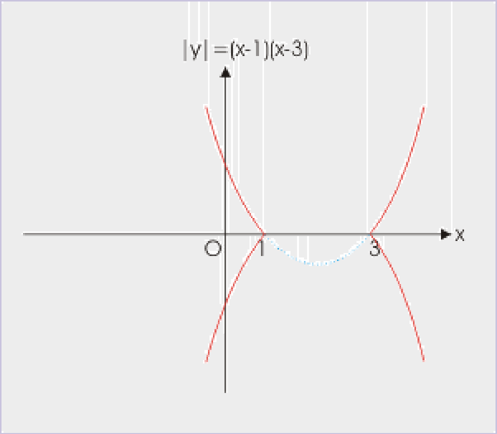
This completes the construction for |y|=f(x).

Example:

Problem : Draw graph of $|y| = (x-1)(x-3)$.

Solution : We first draw the graph of quadratic function given by $y = (x-1)(x-3)$. Then, we remove lower half of the graph and take mirror image of upper half of the graph in x-axis to complete the construction of graph of $|y| = (x-1)(x-3)$.

Modulus operator applied to dependent variable



Modulus operator applied to dependent variable.

Example:

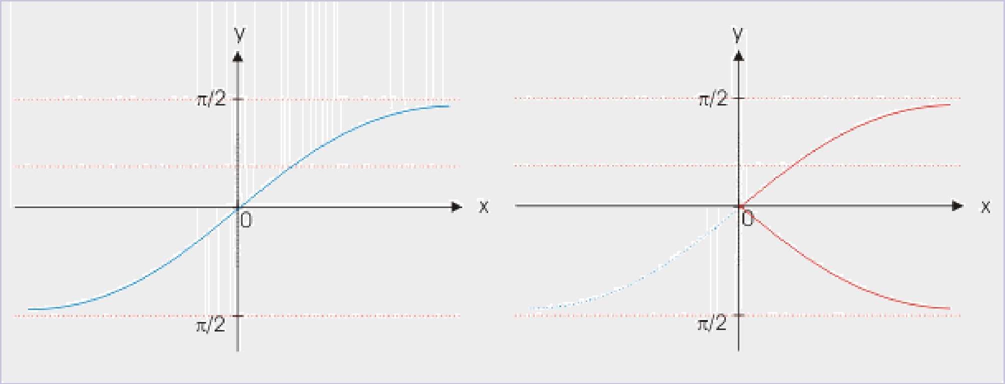
Problem : Draw graph of

$$|y| = \tan^{-1} x$$

.

Solution : We first draw the graph of function given by $y = \tan^{-1} x$. Then, we remove lower half of the graph and take mirror image of upper half of the graph in x-axis to complete the construction of graph of $y = \tan^{-1} x$.

Modulus operator applied to dependent variable



Modulus operator applied to dependent variable.

Modulus function applied to inverse function

The form of transformation is depicted as :

$$y=f(x)\quad\Rightarrow\quad x=|f(y)|$$

The invertible function x= f(y) has its inverse function given by $y=f^{-1}(x)$. Alternatively, if a function is defined as $y=f^{-1}(x)$, then variables x and y are related to each other such that x=f(y). We conclude that graph of $y=f^{-1}(x)$ is same as graph of x=f(y) with the same orientation of x and y axes. It is important to underline here that we transform (change) graph of inverse of given function i.e. $y=f^{-1}(x)$ to get the transformation of graph of x=f(y). Further x and y coordinates on the graph correspond to x and y values.

We interpret assignment of |f(y)| to x in the given graph in accordance with the definition of modulus function. Consider x=|f(y)|. But, modulus can not be equated to negative value. Hence, x can not be negative. It means we need to discard left half of the graph of inverse function $y=f^{-1}(x)$. On the other hand, modulus of negative or positive value is always positive. Hence, positive value of x=a correspond to two values of function in dependent variable, a=±f(y). Corresponding to these two function values in y, we have two values of y i.e. $f^{-1}(a)$ and $f^{-1}(-a)$. In order to plot two values, we need to take mirror image of the left half of the graph of $y=f^{-1}(x)$ across y-axis. This is image in y-axis.

From the point of construction of the graph of x=|f(y)|, we need to modify the graph of $y=f^{-1}(x)$ i.e. x=f(y) as :

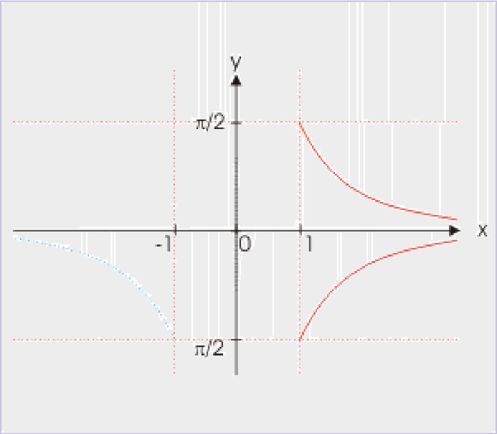
1 : take mirror image of left half of the graph in y-axis

2 : remove left half of the graph

This completes the construction for x=|f(y)|.

Example:
Problem : Draw graph of $x=|\operatorname{cosec} y|;\quad x\in\{-\pi/2,\pi/2\}$.
Solution : The inverse of base function is $\operatorname{cosec}^{-1}x$. We first draw the graph of inverse function. Then, we take mirror image of left half of the graph in y-axis and remove left half of the graph to complete the construction of graph of $x=|\operatorname{cosec} y|$.

Modulus operator applied to function in dependent variable



Modulus operator applied to function in dependent variable.

Examples

Example:

Problem : Find domain of the function given by :

$$f(x) = \frac{1}{\sqrt{|\sin x| + \sin x}}$$

Solution : The square root gives the condition :

$$\Rightarrow |\sin x| + \sin x \geq 0$$

But denominator can not be zero. Hence,

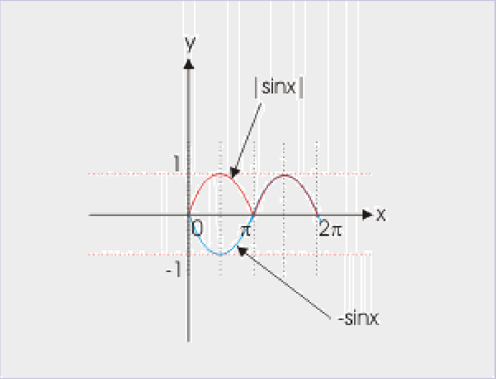
$$\Rightarrow |\sin x| + \sin x > 0$$

$$\Rightarrow |\sin x| > -\sin x$$

We shall make use of graphing technique to evaluate the interval of x. Since both functions are periodic. It would be indicative of the domain if we confine our consideration to 1 period of sine function (0, 2π) and then extend the result subsequently to other periodic intervals.

We first draw sine function. To draw |sinx|, we take image of lower half in x-axis and remove the lower half. To draw “–sinx”, we take image of y=sinx in x-axis.

Domain of function



Domain of function is evaluated by comparing transformed graphs.

From the graph, we see that |sinx| is greater than “-sinx” in (0,π). Note that end points are not included. The domain is written with general notation as :

$$x \in (2n\pi, (2n+1)\pi)$$

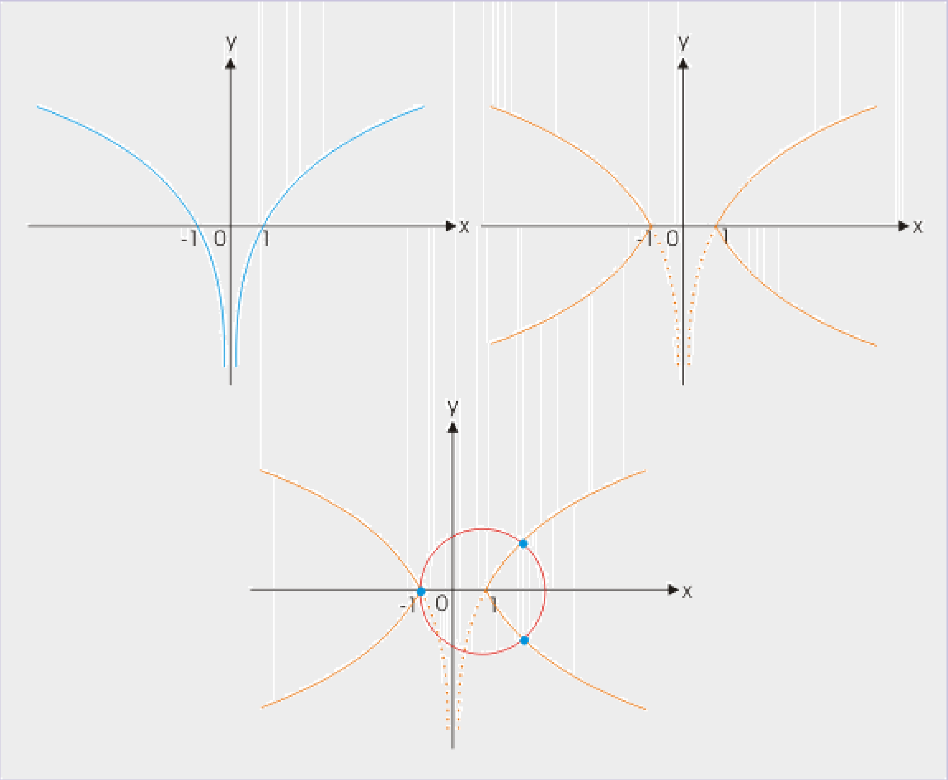
Example:

Problem : Determine graphically the points where graphs of $|y| = \log_e |x|$ and $(x-1)^2 + y^2 - 4 = 0$ intersect each other.

Solution : The function $|y| = \log_e |x|$ is obtained by transforming $y = \log_e x$. To draw $y = \log_e |x|$, we need to remove left half (but here there is no left half) and take image of right half in y-axis. To draw $|y| = \log_e |x|$, we transform the graph of $y = \log_e |x|$. For this, we remove the lower half and take image of upper half in x-axis.

On the other hand, $(x-1)^2 + y^2 - 4 = 0$ is a circle with center at 1,0 having radius of 2 units. Finally, superposing two graphs, we determine the intersection points.

Intersection points



Intersection points are graphically determined.

Clearly, there are three intersection points as shown by solid circles.

Exercises

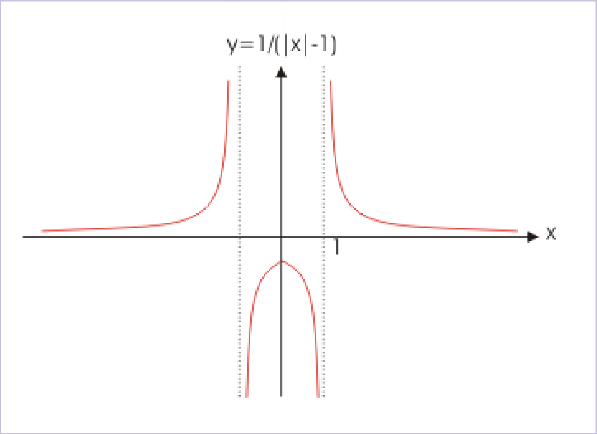
Exercise:

Problem: Draw the graph of function given by :

$$f(x) = \frac{1}{[x] - 1}$$

Hints : Draw 1/x, which is a hyperbola with center at (0,0). Then draw 1/x-1. It is a hyperbola shifted right by 1 unit. Its center is (1,0). Remove left half and take the image of right half in y-axis.

Solution:
Transformation by modulus operator



Transformed graph is shown.

Exercise:

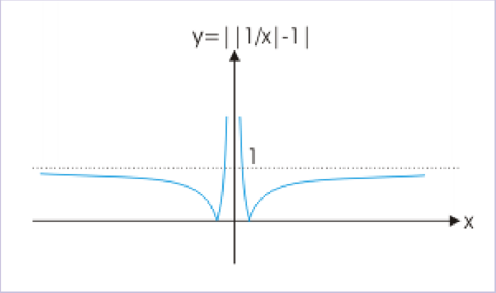
Problem: 2. Draw the graph of function given by :

$$f(x) = \left| \left| \frac{1}{x} \right| - 1 \right|$$

Hints : Draw 1/x, which is a hyperbola with center at (0,0). Then draw |1/x|. Take image of lower half in x-axis. Remove lower half. To draw |1/x|-1, shift down the graph of |1/x| by 1 unit. To draw ||1/x|-1|, Take image of lower half of the graph of |1/x|-1 in x-axis. Remove lower half.

Solution:

Transformation by modulus operator

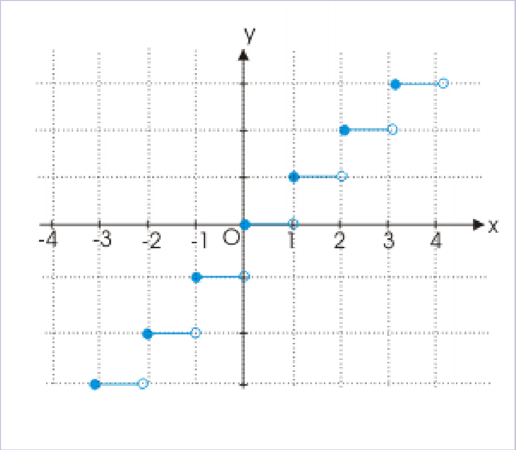


Transformed graph is shown.

Transformation of graphs by greatest integer function

Drawings of graphs resulting from transformation applied by greatest integer function (GIF) follow the same reasoning and steps as deliberated for modulus operator. Before, we proceed to draw graphs for different function forms, we need to recapitulate the graph of greatest integer function (GIF) and also infer thereupon few of the values of GIF around zero.

Graph of GIF



The GIF returns integral values.

For clarity, we apply circular symbols : solid circle to denote inclusion of point and empty circle to denote exclusion of point. Using solid circle is optional, but helps to identify points on the graph. The values of GIF around zero are (we can write these expressions by observing graph. A bit of practice to write down these intervals helps.) :

$$[x] = -3; \quad -3 \le x < -2$$

$$[x] = -2; \quad -2 \le x < -1$$

$$[x] = -1; \quad -1 \le x < 0$$

$$[x] = 0; \quad 0 \le x < 1$$

$$[x] = 1; \quad 1 \le x < 2$$

$$[x] = 2; \quad 2 \le x < 3$$

$$[x] = 3; \quad 3 \le x < 4$$

Important point to note is that lower integer in the interval is included and higher integer is excluded. For negative interval like $-3 \le x < -2$, note that -3 is lower end whereas -2 is higher end. Yet another feature of this function is that domain of the function is continuous in R. It means, there need to be a function value corresponding to all x.

A function like y=f(x) has different elements. We can apply GIF to these elements of the function. There are following different possibilities :

1 : y= f([x])

2 : y=[f(x)]

3 : [y]=f(x)

Greatest Integer operator applied to independent variable

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad y = f([x])$$

The graph of y=f(x) is transformed in y=f([x]) by virtue of changes in the argument values due to operation on independent variable. The independent variable of the function is subjected to greatest function operator. This changes the normal real value input to the function. Instead of real numbers, independent variable to function is rendered to be integers – depending on the value of x and interval it belongs to. A value like x= - 2.3 is passed to the function as -3 in the interval $-3 \le x < -2$.

Clearly, real values of “x” are truncated to integer values in the interval of unity i.e. [-1,0), [0,1), [1.2) etc. It means that values of the function y=f(|x|) will remain same as that of its value corresponding to integral value of “x” till value of “x” changes to next interval. We need to apply modification to the curve to reflect this effect. Knowing that truncation takes place for successive integral values of x, we divide graph of y=f(x) to correspond to 1 unit segments of x-axis. For this, we draw lines parallel to y-axis at integral points along x-axis. From intersection point of lines drawn and function graph, we draw lines parallel to x-axis for the whole interval which extends for a unit value. This ensures that function values remain same to that of function value for the lower integral value of x in a particular interval of one.

From the point of construction of the graph of y=f([x]), we need to modify the graph of y=f(x) as :

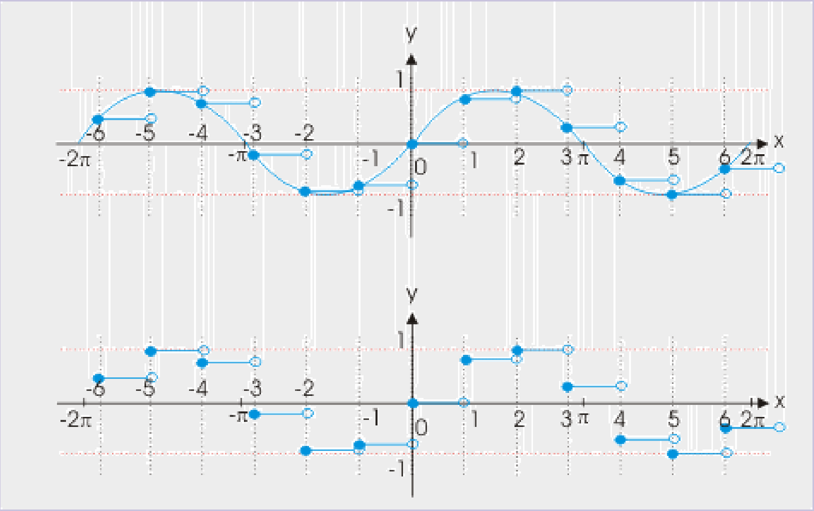
1 : Draw lines parallel to y-axis (vertical lines) at integral values along x-axis to cover the graph of y=f(x).

2 : Identify points of intersections of graph with parallel lines drawn in the earlier step.

3 : Draw lines of 1 unit parallel to x-axis from intersection points in the direction of positive x. The line ends at the next parallel line on right. Include intersection point but exclude other end of the line. Include transformation for all points of the graph.

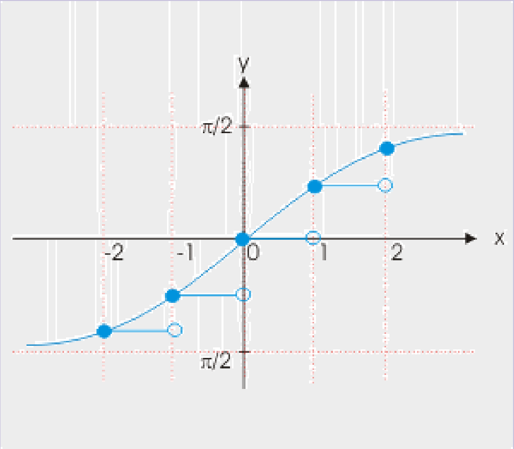
The lines drawn in step 3 is the graph of y=f([x]).

Example:
Problem : Draw the graph of sin[x].
Solution : Following the construction steps, graph of y=sin[x] is drawn as shown here.
Graph of y=sin[x]



The argument of function is modified by GIF.

Example:
Problem : Draw graph of $\tan^{-1}[x]$, $x \in [-2, 2]$.
Solution : Following the construction steps, graph of $y = \tan^{-1}[x]$ is drawn as shown here.
Graph of $y = \tan^{-1}[x]$



The argument of function is modified by GIF.

See that function value corresponding to x=2 and x=-2 are not included in the preceding interval on the graph. As such, we need to put a solid circle at x=2 and x=-2 additionally. Further, we need to remove original graph of y= $\tan^{-1} x$ (this step is not shown in the figure above).

Greatest Integer operator applied to the function

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad y = [f(x)]$$

The graph of y= f(x) is transformed in y=[f(x)] by applying changes to the output of the function. Whatever be the function values, they will be changed to integral values following definition of greatest integer values as given earlier for few intervals. Clearly, real values of “f(x)” are truncated to integer values in the interval of unity i.e. [-1,0), [0,1), [1.2) etc along y-axis.

From the point of construction of the graph of y=f([x]), we need to modify the graph of y=f(x) as :

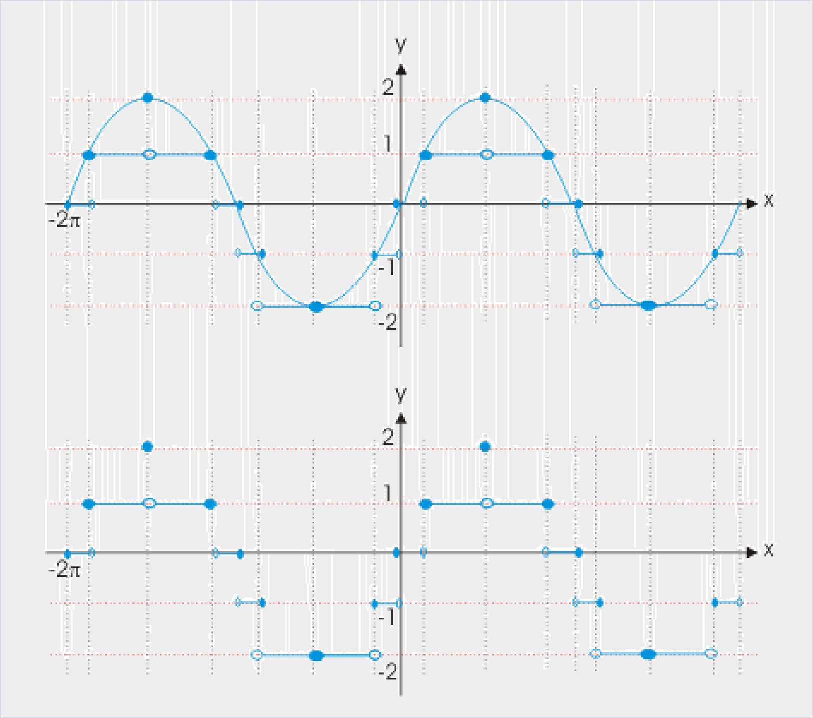
1 : Draw lines parallel to x-axis (horizontal lines) at integral values along y-axis to cover the graph of y=f(x).

2 : Identify points of intersections of graph with parallel lines drawn in the earlier step. Draw lines parallel to y-axis (vertical lines) from the intersection points identified.

3 : Take x-projection of curve from the point of intersection between two consecutive vertical lines such that it lies on horizontal line of lower value. Include intersection point but exclude other end of the line. Further include points not covered by the projection.

The lines drawn in step 3 is the graph of y=[f(x)].

Example:
Problem : Draw the graph of [2sinx].
Solution : Following the construction steps, graph of y=[2sinx] is drawn as shown here.
Graph of y=[2sinx]



The value of function is modified by GIF.

Values assigned to greatest Integer function

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad [y] = f(x)$$

We need to evaluate this equation on the basis of assignment to the dependent expression. The value of function f(x) is first calculated for a given value of x. The value so evaluated is assigned to the GIF function [y]. We interpret assignment to [y] in accordance with the interpretation of equality of the GIF function to a value. In this case, we know that :

$$[y] = f(x); \quad f(x) \notin Z \quad \Rightarrow \text{GIF can not be equated to non-integers. No solution.}$$

$$[y] = f(x); \quad f(x) \in Z \quad \Rightarrow \text{y} = \text{Continuous interval of 1 unit starting from f(x)}$$

Clearly, we need to neglect plot corresponding to all non-integral values of f(x). For every value of x, which yields integral value of f(x), there are multiple values of dependent expression [y] in an interval of 1 unit. For example, for $[y] = f(x) = 2,y \in 2 \le y < 3$. In the nutshell, this graph is not continuous. There is no value of y corresponding to non integer f(x) and there are multiple values of y in an interval of 1 for integral values of f(x).

From the point of construction of the graph of |y|=f(x), we need to modify the graph of y=f(x) as :

1 : Draw lines parallel to x-axis (horizontal lines) at integral values along y-axis to cover the graph of y=f(x).

2 : Identify points of intersections of graph with parallel lines (horizontal lines) drawn in the earlier step.

3 : Draw lines of 1 unit parallel to y-axis (vertical lines) from intersection points in the positive y-direction. Include intersection point but exclude other end of the line.

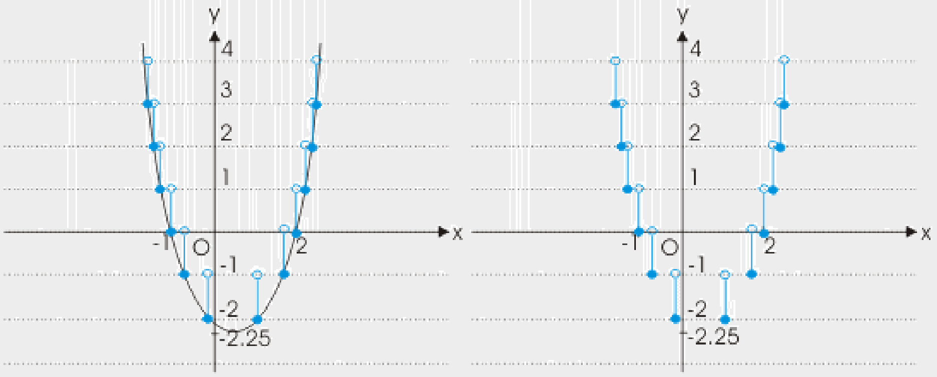
The lines drawn in step 3 is the graph of [y]= f(x).

Example:
Problem : Draw graph of [y]=(x+1)(x-2).
Solution : We first draw the graph of quadratic polynomial function $y = (x+1)(x-2) = x^2 - x - 2$. The lowest point of the parabola is calculated as :

$$D = (-1)^2 - (4X1X - 2) = 1 + 8 = 9$$

$$y_{\text{min}} = -\frac{D}{4a} = -\frac{9}{4X1} = -2.25$$

Following construction steps, graph of [y]=(x+1)(x-2) is drawn as shown here.
Graph of [y]=(x+1)(x-2)

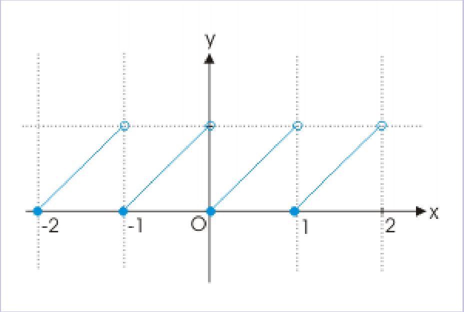


The value of expression is assigned to GIF.

Transformation applied by fraction part function

Drawings of graphs resulting from transformation applied by fraction part function (FPF) follow the same reasoning and steps as deliberated for modulus and greatest integer function. Before, we proceed to draw graphs for different function forms, we need to recapitulate the graph of fraction part function (FPF) and also infer thereupon few of the values of FPF around zero.

Fraction part function



FPF is a periodic function with period 1.

The values of FPF around zero are (we can write these expressions by observing graph. A bit of practice to write down these intervals helps.) :

$$\{x\} = x + 2; \quad -2 \le x < -1$$

$$\{x\} = x + 1; \quad -1 \le x < 0$$

$$\{x\} = x; \quad 0 \le x < 1$$

$$\{x\} = x - 1; \quad 1 \le x < 2$$

Important point to note is that lower integer is included, but higher integer is excluded in the intervals of unity in which FPF is defined. The graph segment in the interval [0,1) is y=x i.e. identity function. We obtain expression of function in right intervals (positive value intervals of x) by shifting identity function towards right by 1 successively and in left intervals (negative value intervals of x) by shifting identity function towards left by 1 successively. The values of FPF are continuous real values which is equal to or greater than zero but less than 1. These function values are repeated in each of intervals of unity along x-axis. Thus, FPF is a periodic function with period of 1. Domain of FPF is R and range is [0,1). Further, FPF is related to real number as x=[x]+{x}.

A function like y=f(x) has different elements. We can apply FPF to these elements of the function. There are following different possibilities :

1 : y = f({x})

2 : y ={f(x)}

3 : {y} = f(x)

Fraction part operator applied to the argument

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad y = f(\{x\})$$

The graph of y=f(x) is transformed in y=f({x}) by virtue of changes in the argument values. The independent variable is subjected to fraction part operator. This changes the normal real value input to function. Instead of real numbers, independent variable to function is rendered to be fractions irrespective of values of x. A value like x = - 2.3 is passed to the function as 0.7 in the interval [0,1).

Clearly, real values of “x” are truncated to fraction values in all intervals. It means that same set of values of the function y=f(|x|) corresponding to interval of x defined by [0,1] will repeat in other intervals along x-axis. The FPF is a periodic function with a period of 1. Taking advantage of this fact, we obtain graph of y=f({x}) by repeating part of graph for x in [0,1) to other intervals along x-axis. Clearly, transformed function y=f({x}) is periodic with a period of 1.

From the point of construction of the graph of y=f({x}), we need to modify the graph of y=f(x) as :

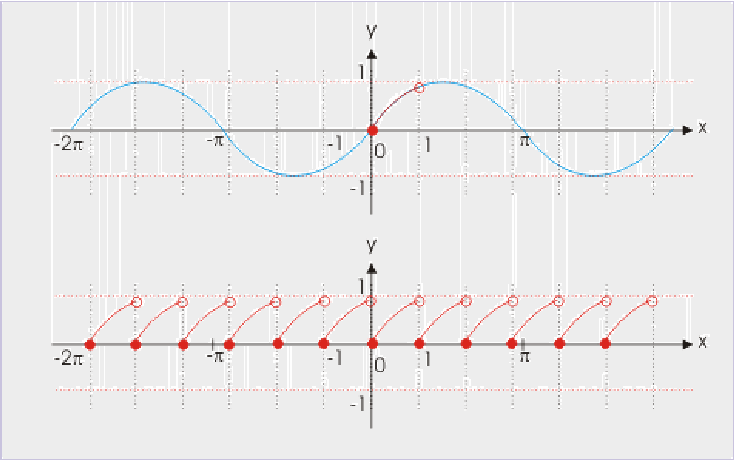
1 : Draw lines parallel to y-axis (vertical lines) at integral values along x-axis to cover the graph of y=f(x).

2 : Identify part of the graph for values of x in [0,1). Include end point corresponding to x=0 and exclude end point corresponding to x=1.

3 : Repeat the part of the graph identified in step 2 for other intervals of x

The lines drawn in step 3 is the graph of y=f({x}).

Example:
Problem : Draw the graph of sin{x}.
Solution : Following the construction steps, graph of y=sin{x} is drawn by transforming y = sinx as shown here.
Graph of y=sin{x}



Repeat the part of the graph identified between 0 and 1 to other intervals of x.

Example:

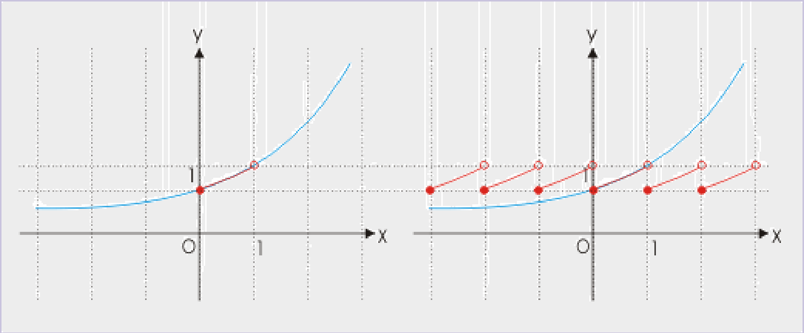
Problem : Draw the graph of $y = \frac{e^x}{e^{[x]}}$.

Solution : Rearranging, we have :

$$\Rightarrow y = e^x - e^{[x]} = e^{x-[x]} = e^{\{x\}}$$

Following the construction steps, graph of $y = e^{\{x\}}$ is drawn by transforming $y = e^x$ as shown here.

Transformation of exponential graph



Repeat the part of the graph identified between 0 and 1 to other intervals of x.

Fraction part function applied to the function

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad y = \{f(x)\}$$

The graph of y= f(x) is transformed in y={f(x)} by applying changes to the output of the function. Whatever be the function values, they will be changed to fraction values following definition of fraction part values as given earlier for few intervals. The values of y will lie in the interval [0,1).

Here, we need to recognize one important aspect of graph of real valued function. Consider a function value y=3. The function value such as y=3.3 shows a change in function value of 3.3-3=0.3. This change in function value depends on the integral part of y, which is 3. The change will be different at other integral part like 2 depending on the nature of function y = f(x). What it means that the nature of graph in the integral intervals of y have different set of fractional parts. In turn it means that when real values are converted to fractional part, resulting values represent different set of fraction parts, which is represented by the nature of graph segment between two consecutive integral intervals of y. Mathematically,

$$\{y\} = y - [y]$$

Clearly, {y} depends on y, but lies in the interval of y given by [0,1).

From the point of construction of the graph of y={f(x)}, we need to modify the graph of y=f(x) as :

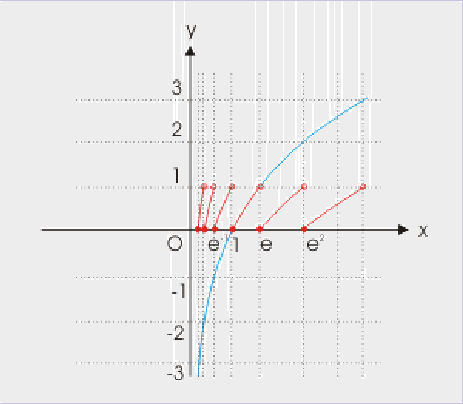
1 : Draw lines parallel to x-axis (horizontal lines) at integral values along y-axis to cover the graph of y=f(x).

2 : Identify segments of graph between two consecutive vertical intervals. Transfer these segments to y interval given by [0,1).

3 : Include end point corresponding to y=0 and exclude end point corresponding to y=1.

The lines drawn in step 3 is the graph of y={f(x)}.

Example:
Problem : Draw the graph of $\{\log_e x\}$.
Solution : Following the construction steps, graph of $y = \{\log_e x\}$ is drawn as shown here.
Transformation of sine graph



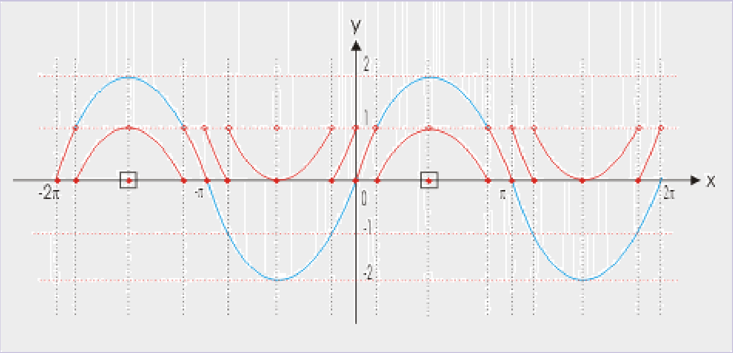
Transfer part of the graph identified in unit y interval [0,1).

Example:

Problem : Draw the graph of {2sinx}.

Solution : Following the construction steps, graph of y={2sinx} is drawn by transforming y= 2sinx as shown here.

Graph of y={sinx}



Transfer part of the graph identified in unit y-interval [0,1).

Note two individual solid circles on x-axis. They have been enclosed in squares for emphasis. We should analyze their existence while constructing the graph.

Values assigned to fraction part function

The form of transformation is depicted as :

$$y = f(x) \quad \Rightarrow \quad \{y\} = f(x)$$

We need to evaluate this equation on the basis of assignment to the dependent expression variable. The value so evaluated is assigned to the FPF function {y}. We interpret assignment to {y} in accordance with the interpretation of equality of the FPF function to a value. In this case, we know that :

$$\{y\} = f(x); \quad f(x) \in Z \quad \Rightarrow \text{FPF can not be equated to integers. No solution.}$$

$$\{y\} = f(x); \quad f(x) \notin Z \quad \Rightarrow \text{y} = \text{Continuous interval of fraction values starting from f(x)}$$

Clearly, we need to neglect plot corresponding to integral values of f(x). On the other hand, there are multiple non-integral values of f(x) for a particular value of x corresponding to different intervals of unity along y. For example,

{-1.47}= {-0.47}= {0.53} = {1.53) = (2.53) = ….. = 0.53

Such is the case with other fractional values. It means that part of the graph of y=f(x) lying in y interval of [0,1) will be repeated in consecutive intervals of 1 along y-axis.

From the point of construction of the graph of {y}= f(x), we need to modify the graph of y=f(x) as :

1 : Draw lines parallel to x-axis (horizontal lines) at integral values along y-axis to cover the graph of y=f(x).

2 : Identify part of the graph in y interval [0,1). Include end point corresponding to y=0 and exclude end point corresponding to y=1. Neglect other part of graph.

3 : Repeat part of graph identified in step 2 in other y intervals of unity along y-axis.

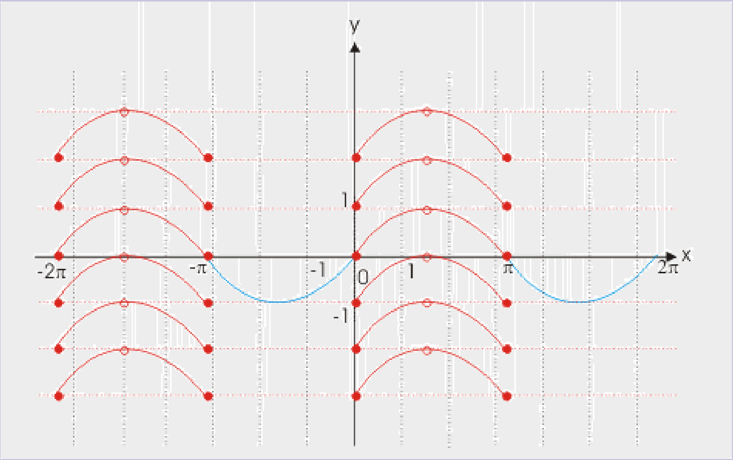
The lines drawn in step 3 is the graph of {y}= f(x).

Example:

Problem : Draw graph of {y}= sinx; x∈[-2π,2π].

Solution : Following construction steps, graph of {y}= sinx is drawn by transforming y= sinx as shown here.

Graph of y=sin{x}



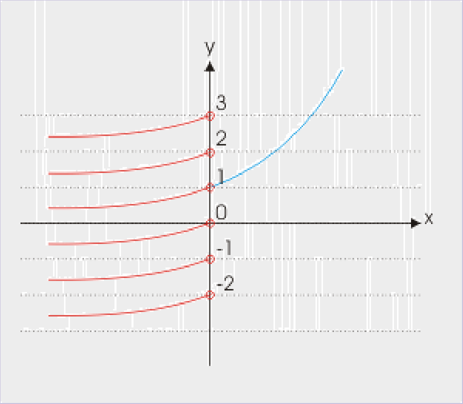
Identify part of the graph in y interval [0,1). Repeat part of graph so identified in other y intervals of unity

Example:

Problem : Draw graph of $\{y\} = e^x$.

Solution : Following construction steps, graph of $\{y\} = e^x$ is drawn by transforming $y = e^x$ as shown here.

Graph of y=sin{x}



Identify part of the graph in y interval [0,1). Repeat part of graph so identified in other y intervals of unity

Equal functions

Two numbers are equal if they are same number. Two variables are equal if they represent same number. Following these connotations, two functions are equal if they are same function. But, very concept of “equal” or “identical” functions indicates that there is more than one way to represent a function. In other words, the question of equality of two functions arises when two function forms yield same values. There are few such occurrences in mathematics. This arises primarily because we have alternate ways to represent a mathematical entity. Consider, for example, modulus function. There are two equivalent expressions :

$$f(x) = |x|$$

$$g(x) = \sqrt{x^2}$$

These two function forms yield same values for all real values of x. As such, these two functions f(x) and g(x) are equal functions. On the other hand, there are equivalent forms, which represent equal values but not for all values of x in the domains of two definition. Consider, for example,

$$f(x) = 2\log_e x$$

$$g(x) = \log_e x^2$$

The logarithmic function f(x) is defined for x>0. This means its domain is (0, ∞). For logarithmic function, g(x),

$$\Rightarrow x^2 > 0$$

This inequality is true for all values of x except x=0. It means domain of g(x) is R-{0}. Clearly, domains of two functions are not equal. For a value x = -1, g(x) yields a value while f(x) is not defined for this value of x. Two equations, therefore, are not equal. However, two functions are equal if we limit our consideration for domain limited to the intersection of two domains. Hence,

$$f(x) = g(x); \quad x \in (0,\infty)$$

There is yet another possibility. Two equivalents forms have same domains, but yield different set of values. In such case also, two functions are not equal. Consider the example given here.

Example:
Problem : Determine whether f(x) and g(x) are identical functions?

$$f(x) = x$$

$$g(x) = \sqrt{x^2}$$

Solution : Here, f(x) is defined for all values of x and its domain is R. On the other hand, domain of g(x) is also R as square of x is always non-negative. However, square root of a number is non-negative. Therefore, two function forms are not equivalent as f(x) is real, whereas is g(x) is non-negative and is a subset of R. Thus, range of f(x) is R and range of g(x) is (0,∞). Clearly, two given functions are not equal.

In the nutshell, two equivalent function forms are equal if their domain, range and function values are equal.

Definition of equal functions

Two functions f(x) and g(x) are equal functions, if :

(i) Domain of f (x) = Domain of g(x) = X

(ii) f(x) = g(x) for all $x \in X$

Equal functions are also known as identical functions. Above two conditions are sufficient for two functions to be equal. Since second condition means that values of functions are equal for every x in the domain, it is guaranteed that range of two functions are equal.

Range of f (x) = Range of g(x) = Y

Examples

Example:
Problem : Determine whether f(x) and g(x) are identical functions.

$$f(x) = \frac{x}{x^2}$$

$$g(x) = \frac{1}{x}$$

Solution : Two function forms are equivalent as f(x) is reduced to g(x) on simplification. Now, expression of f(x) is defined for all values of x except x=0. Thus, domain of f(x) is R-{0}. On the other hand, domain of reciprocal function g(x) is also R-{0}. Clearly, two given functions are equal.

Example:
Problem : 3. Determine whether f(x) and g(x) are identical functions.

$$f(x) = \log_e x - \log_e\left(x^2+1\right)$$

$$g(x) = \log{}_e\left(\frac{x}{1+x^2}\right)$$

Solution :
Two function forms are equivalent as f(x) is changed to g(x) and vice-versa on simplification. Now, f(x) is defined for

$$x > 0 \quad \text{and} \quad x^2+1 > 0$$

But x^2 is always positive. Hence, domain of f(x) is (0, ∞). On the other hand, g(x) is defined for :

$$\frac{x}{1+x^2} > 0$$

$$x > 0$$

Thus, domain of g(x) is also (0, ∞). Hence, two functions are identical.

Example:

Problem : Determine domains for which two functions are equal.

$$f(x) = \log x - \log(x-1)$$

$$g(x) = \log\left(\frac{x}{x-1}\right)$$

Solution : Two function forms are equivalent as f(x) is changed to g(x) and vice-versa on simplification. Now, f(x) is defined for

$$x > 0 \quad \text{and} \quad x - 1 > 0$$

$$x > 0 \quad \text{and} \quad x > 1$$

Hence, domain of f(x) is intersection of two intervals (1, ∞). On the other hand, g(x) is defined for :

$$\frac{x}{x-1} > 0$$

Critical points are 0 and 1. Using sign rule for rational function, the domain of g(x) is values of x satisfying above inequality :

$$(-\infty, 0) \cup (1,\infty)$$

Clearly, two domains are not equal. Note that there is no restriction on the range of the functions. Therefore, two functions are equal in the restricted domain which is intersection of two domains.

$$\text{Domain} = (1,\infty)$$

Acknowledgment

Author wishes to thank Ms. Aditi Singh, New Delhi for her valuable suggestions on the topic.

Even and odd functions

Even and odd functions are related to symmetry of functions. The symmetry of a function is visualized by the planar plot of a function, which may show symmetry with respect to either an axis (y-axis) or origin.

Since functions need not always be symmetric, they may neither be even nor be odd. The parity of a function i.e. whether it is even or odd is determined with certain algebraic algorithm. Further, symmetry of functions may change subsequent to mathematical operations.

Even functions

The values of even function at x=x and x=-x are same.

Even function

> A function f(x) is said to be "even" if for every "x", there exists "-x" in the domain of the function such that :

$$f(-x) = f(x)$$

An even function is symmetric about y-axis. If we consider the axis as a mirror, then the plot in first quadrant has its mirror image (bilaterally inverted) in second quadrant. Similarly, the plot in fourth quadrant has its mirror image (bilaterally inverted) in third quadrant.

Some examples of even functions are $x^2, \ x$ and $\cos x$. In each case, we see that :

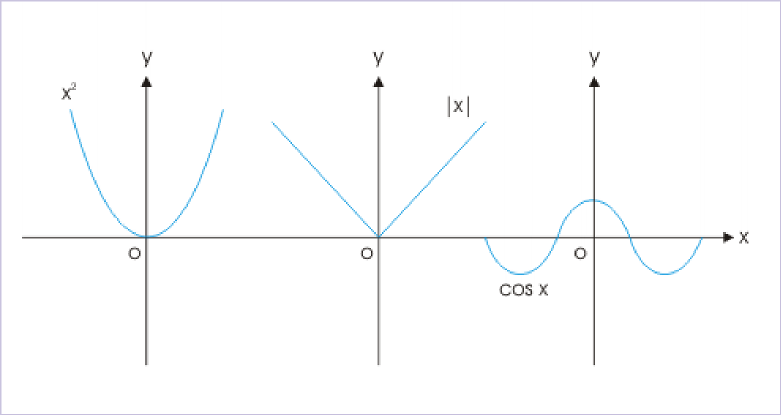
$$\Rightarrow f(-x) = (-x)^2 = x^2 = f(x)$$

$$\Rightarrow f(-x) = |-x| = |x| = f(x)$$

$$\Rightarrow f(-x) = \cos(-x) = \cos x = f(x)$$

The right side is mirror image of left hand side and the left side is mirror image of right hand side of the curve.

Even functions



Examples of even functions.

It is important to see that if we rotate the curve by 180° about y-axis, then the appearance of the rotated curve is same as the original curve. We can state this alternatively as : if we rotate left hand side of the curve by 180° about y-axis, then we get the right hand curve and vice-versa.

Examples

Problem 1: Prove that the function f(x) is "even", if

$$f(x) = x\frac{a^x - 1}{a^x + 1}$$

Solution : For function being "even", we need to prove that :

$$f(-x) = f(x)$$

Here,

$$\Rightarrow f(-x) = -x\frac{a^{-x} - 1}{a^{-x} + 1} = -x\frac{\frac{1}{a^x} - 1}{\frac{1}{a^x} + 1}$$

$$\Rightarrow f(-x) = -x\frac{\frac{1-a^x}{a^x}}{\frac{1+a^x}{a^x}} = -x\frac{1 - a^x}{1 + a^x}$$

$$\Rightarrow f(-x) = x\frac{a^x - 1}{a^x + 1} = f(x)$$

Problem 2: If an even function "f" is defined on the interval (-5,5), then find the real values for which

$$f(x) = f\left(\frac{x+1}{x+2}\right)$$

Solution : It is given that function "f" is even. Hence, arguments of the functions on two sides are related either as

$$\Rightarrow x = \frac{x+1}{x+2}$$

or as :

$$\Rightarrow x = -\frac{x+1}{x+2}$$

From the first relation,

$$\Rightarrow x^2 + x - 1 = 0$$

$$\Rightarrow x = \frac{-1 \pm \sqrt{5}}{2}$$

From the second relation,

$$\Rightarrow x^2 + 3x + 1 = 0$$

$$\Rightarrow x = \frac{-3 \pm \sqrt{5}}{2}$$

We see that values are within the specified domain. Hence, all the four solutions satisfy the given equation.

Odd functions

The values of odd function at x=x and x=-x are equal in magnitude but opposite in sign.

Odd function

> A function f(x) is said to be "odd" if for every "x", there exists "-x" in the domain of the function such that :

$$f(-x) = -f(x)$$

An odd function is symmetric about origin of the coordinate system. The plot in first quadrant has its mirror image (bilaterally inverted) in third quadrant. Similarly, the plot in second quadrant has its mirror image (bilaterally inverted) in fourth quadrant.

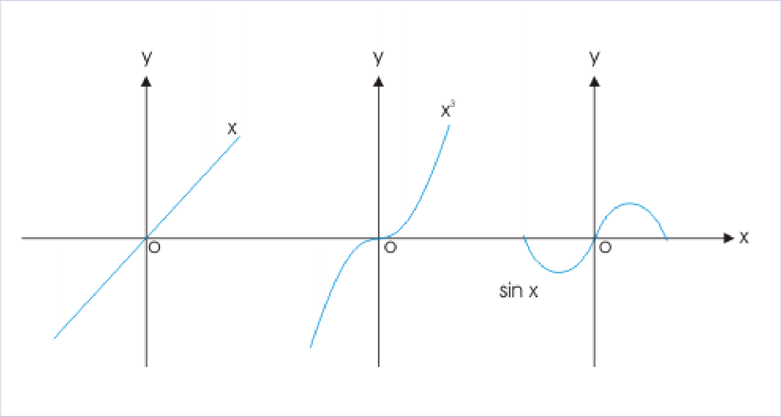
Some examples of odd functions are :$x, x^3 \quad \text{and} \quad \sin x$. In each case, we see that :

$$\Rightarrow f(-x) = -x = -f(x)$$

$$\Rightarrow f(-x) = (-x)^3 = -x^3 = -f(x)$$

$$\Rightarrow f(-x) = \sin(-x) = -\sin x = -f(x)$$

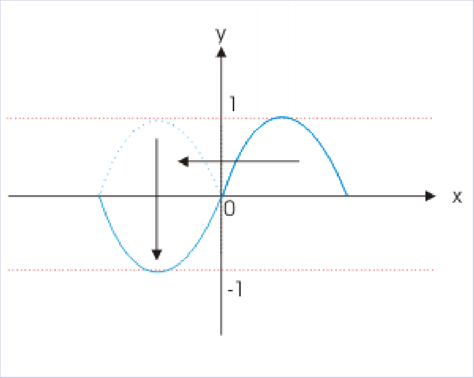
The upper curve of these functions is exactly same as the lower curve across x-axis.
Odd functions



Examples of odd functions.

It is important to see that if we rotate the curve by 180° about origin, then the appearance of the rotated curve is same as the original curve. In other words, if we rotate right hand side of curve by 180° about origin, then we get left side of the curve. Further, it is interesting to note that we obtain left hand part of the plot of odd function in two steps : (i) drawing reflection (mirror image) of right hand plot about y-axis and (ii) drawing reflection (mirror image) of "reflection drawn in step 1" about x-axis.

Odd function plot



Odd function as two successive mirror images

Examples

Problem 3: Determine whether the function f(x) is "odd" function, where :

$$f(x) = \log_e \left\{ x + \sqrt{(x^2 + 1)} \right\}$$

Solution : In order to determine the nature of function with respect to even or odd, we check for f(-x). Here,

$$\Rightarrow f(-x) = \log_e\left[-x + \sqrt{\left\{(-x)^2 + 1\right\}}\right] = \log_e\left\{-x + \sqrt{(x^2+1)}\right\}$$

The expression on the right hand side can not be explicitly interpreted whether it equals to f(x) or not. Therefore, we rationalize the expression of logarithmic function,

$$\Rightarrow f(-x) = \log_e \frac{\left\{-x + \sqrt{(x^2+1)}\right\} X \left\{x + \sqrt{(x^2+1)}\right\}}{\left\{x + \sqrt{(x^2+1)}\right\}} = \log_e \frac{-x^2 + x^2 + 1}{\left\{x + \sqrt{(x^2+1)}\right\}}$$

$$\Rightarrow f(-x) = \log_e 1 - \log_e\left\{x + \sqrt{(x^2+1)}\right\} = -\log_e\left\{x + \sqrt{(x^2+1)}\right\} = -f(x)$$

Hence, given function is an “odd” function.

Problem 4: Determine whether sinx + cosx is an even or odd function?

Solution : In order to check the nature of the function, we evaluate f(-x),

$$f(-x) = \sin(-x) + \cos(-x) = -\sin x + \cos x$$

The resulting function is neither equal to f(x) nor equal to “-f(x)”. Hence, the given function is neither an even nor an odd function.

Mathematical operations and nature of function

It is easy to find the nature of function resulting from mathematical operations, provided we know the nature of operand functions. As already discussed, we check for following possibilities :

- If f(-x) = f(x), then f(x) is even.
- If f(-x) = -f(x), then f(x) is odd.
- If above conditions are not met, then f(x) is neither even nor odd.

Based on above algorithm, we can determine the nature of resulting function. For example, let us determine the nature of "fog" function when “f” is an even and “g” is an odd function. By definition,

$$fog(-x) = f(g(-x))$$

But, “g” is an odd function. Hence,

$$\Rightarrow g(-x) = -g(x)$$

Combining two equations,

$$\Rightarrow fog(-x) = f(-g(x))$$

It is given that “f” is even function. Therefore, f(-x) = f(x). Hence,

$$\Rightarrow fog(-x) == f(-g(x)) = f(g(x)) = fog(x)$$

Therefore, resulting “fog” function is even function.

The nature of resulting function subsequent to various mathematical operations is tabulated here for reference :

```
----------------------------------------------------------------
-------------------- f(x) g(x) f(x) ± g(x) f(x) g(x) f(x)/g(x),
g(x)≠0 fog(x) --------------------------------------------------
---------------------------------- odd odd odd even even odd odd
even Neither odd odd even even even even even even even --------
----------------------------------------------------------------
------------
```

We should emphasize here that we need not memorize this table. We can always carry out particular operation and determine whether a particular operation results in even, odd or neither of two function types. We shall work with a division operation here to illustrate the point. Let f(x) and g(x) be even and odd functions respectively. Let h(x) = f(x)/g(x). We now substitute “x” by “-x”,

$$\Rightarrow h(-x) = \frac{f(-x)}{g(-x)}$$

But f(x) is an even function. Hence, f(-x) = f(x). Further as g(x) is an odd function, g(-x) = -g(x).

$$\Rightarrow h(-x) = \frac{f(x)}{-g(x)} = -h(x)$$

Thus, the division, here, results in an odd function.

There is an useful parallel here to remember the results of multiplication and division operations. If we consider even as "plus (+)" and odd as "minus (-)", then the resulting function is same as that resulting from multiplication or division of plus and minus numbers. Product of even (plus) and odd (minus) is minus(odd). Product of odd (minus) and odd (minus) is plus (even). Similarly, division of odd (minus) by even (plus) is minus (odd) and so on.

Square of an even or odd function

The square of even or odd function is always an even function.

Properties of derivatives

1: If f(x) is an even differentiable function on R, then f’(x) is an odd function. In other words, if f(x) is an even function, then its first derivative with respect to "x" is an odd function.

2: If f(x) is an odd differentiable function on R, then f’(x) is an even function. In other words, if f(x) is an odd function, then its first derivative with respect to "x" is an even function.

Composition of a function

Every real function can be considered to be composed from addition of an even and an odd function. This composition is unique for every real function. We follow an algorithm to prove this as :

Let f(x) be a real function for x ∈ R. Then,

$$f(x) = f(x) + f(-x)\} - f(-x)\}$$

Rearranging,

$$f(x) = \frac{1}{2}\left\{f(x) + f(-x)\right\} + \frac{1}{2}\left\{f(x) - f(-x)\right\} = g(x) + h(x)$$

Now, we seek to determine the nature of functions “g(x)” and “h(x). For “g(x)”, we have :

$$\Rightarrow g(-x) = \frac{1}{2}\left[f(-x) + f\left\{-(-x)\right\}\right] = \frac{1}{2}\left\{f(-x) + f(x)\right\} = g(x)$$

Thus, “g(x)” is an even function.

Similarly,

$$\Rightarrow h(-x) = \frac{1}{2}\left[f(-x) - f\left\{-(-x)\right\}\right] = \frac{1}{2}\left\{f(-x) - f(x)\right\} = -h(x)$$

Clearly, “h(x)” is an odd function. We, therefore, conclude that all real functions can be expressed as addition of even and odd functions.

Even and odd extensions of function

A function has three components – definition(rule), domain and range. What could be the meaning of extension of function? As a matter of fact, we can not extend these components. The concept of extending of function is actually not a general concept, but limited with respect to certain property of a function. Here, we shall consider few even and odd extensions. Idea is to complete a function defined in one half of its representation (x>=0) with other half such that resulting function is either even or odd function.

Even function

Let f(x) is defined in [0,a]. Then, even extension is defined as :

`|f(x); 0≤x≤a g(x) = | | f(-x); -a≤x<0`

The graphical interpretation of such extension is that graph of function f(x) is extended in other half which is mirror image of f(x) in y-axis i.e. image across y-axis.

Odd extension

Let f(x) is defined in [0,a]. Then, odd extension is defined as :

`| f(x); 0≤x≤a g(x) = | | -f(x); -a≤x<0`

The graphical interpretation of such extension is that graph of function f(x) is extended in other half which is mirror image of f(x) in x-axis i.e. image across x-axis.

Exercises

Exercise:

Problem: Determine whether f(x) is odd or even, when :

$$f(x) = e^{x} + e^{-x}$$

Solution:

The function “f(x)” consists of exponential terms. Here,

$$\Rightarrow f(-x) = e^{-x} + e^{-(-x)} = e^{-x} + e^{x} = e^{x} + e^{-x} = f(x)$$

Hence, given function is even function.

Exercise:

Problem: Determine whether f(x) is odd or even, when :

$$f(x) = \frac{x}{e^{x} - 1} + \frac{x}{2}$$

Solution:

The function “f(x)” consists of exponential terms. In order to check polarity, we determine f(-x) :

$$f(-x) = -\frac{x}{e^{-x}-1} + \frac{-x}{2} = -\frac{x}{1/e^{x}-1} - \frac{x}{2}$$

$$\Rightarrow f(-x) = -\frac{xe^{x}}{1-e^{x}} - \frac{x}{2}$$

We observe here that it might be tedious to reduce the expression to either “f(x)” or “-f(x)”. However, if we evaluate f(x) – f(-x), then the resulting expression can be easily reduced to simpler form.

$$f(x) - f(-x) = \frac{x}{e^{x}-1} + \frac{x}{2} + \frac{xe^{x}}{1-e^{x}} + \frac{x}{2}$$

$$\Rightarrow f(x) - f(-x) = \frac{x}{e^{x}-1} - \frac{xe^{x}}{e^{x}-1} + x = \frac{x(1-e^{x})}{e^{x}-1} + x = 0$$

Hence,

$$f(x) = f(-x)$$

It means that given function is an even function.

Exercise:

Problem:) How to check whether a pulse equation of the form

$$y = \frac{a}{\left\{(3x+4t)^{2}+b\right\}}$$

is symmetric or asymmetric, here "a" and "b" are constants.

Note: Posted by Dr. R.K.Singhal through e-mail

Solution:

The pulse function has two independent variables “x” and “t”. The function needs to be even for being symmetric about y-axis at a given instant, say t =0.

We check the nature of function at t = 0.

$$\Rightarrow y = \frac{a}{(9x^{2}+b)}$$

$$\Rightarrow f(-x) = \frac{a}{\left\{9(-x)^2 + b\right\}} = \frac{a}{(9x^2 + b)} = f(x)$$

Thus, we conclude that given pulse function is symmetric.

Exercise:

Problem: Determine whether f(x) is odd or even, when :

$$f(x) = x^2 \cos x - \ \ \sin x$$

Solution:

The “f(x)” function consists of trigonometric and modulus functions. Here,

$$\Rightarrow f(-x) = (-x)^2 \cos(-x) - \ \ \sin(-x)$$

We know that :

$$(-x)^2 = x^2; \quad \cos(-x) = \cos x; \qquad \sin(-x) \ = \ -\sin x \ = \ \sin x$$

Putting these values in the expression of f(-x), we have :

$$\Rightarrow f(-x) = (-x)^2 \cos(-x) - \ \ \sin(-x) \ = x^2 \cos x - \ \ \sin x \ = f(x)$$

Hence, given function is an even function.

Exercise:

Problem: Determine whether f(x) is odd or even, when :

$$f(x) = xe^{-x^2 \tan^2 x}$$

Solution:

The “f(x)” function consists of exponential terms having trigonometric function in the exponent. Here,

$$\Rightarrow f(-x) = (-x)e^{-\left\{(-x)^2 \tan^2(-x)\right\}}$$

We know that :

$$(-x)^2 = x^2; \quad \tan^2(-x) = (-\tan x)^2 = \tan^2 x$$

$$\Rightarrow f(-x) = (-x)e^{-\left\{(-x)^2 \tan^2(-x)\right\}} = -xe^{-x^2 \tan^2 x} = -f(x)$$

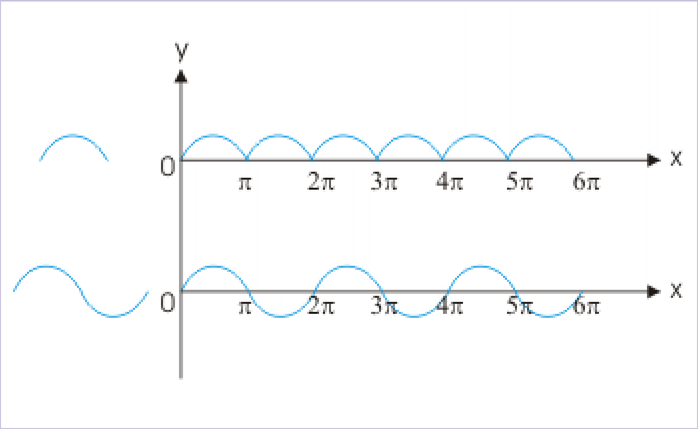
Hence, given function is an odd function.

Periodic functions

In physical world around us, we encounter many phenomena which repeat after certain interval of time. In mathematics, the notion of periodicity remains same but with more general connotation. The periodicity of a function is not limited to time. We look for repetition of function values with respect to independent variable. Time could be just one such independent variable. For example, we have seen that trigonometric functions are “many one” relations. This means that we get same value of trigonometric function for different angles. This “many one” relation is the basic requirement for a function to be periodic. In addition, these same values of the function should appear at regular intervals for the values of independent variables in the domain.

We can visualize periodic nature of a function by observing its graph in which a particular smallest segment of the plot can be repeated to construct complete plot.

Periodic function



A sine curve is constructed by repeating a segment of the curve as shown in the lower graph in the figure.

In the two graphs shown above, we have considered two segments corresponding to intervals of “π” and “2π”. The graphs are constructed by repeating the segments one after another. It is clear from the figure that we need smallest segment of an interval “2π” to construct a sine curve (lower curve).

Definition of periodic function

A function is said to be periodic if there exists a positive real number “T” such that

$$f(x+T) = f(x) \quad \text{for all} \quad x \in D$$

where “D” is the domain of the function f(x). The least positive real number “T” (T>0) is known as the fundamental period or simply the period of the function. The “T” is not a unique positive number. All integral multiple of “T” within the domain of the function is also the period of the function. Hence,

$$f(x+nT) = f(x); \quad n \in Z, \quad \text{for all} \quad x \in D$$

In the context of periodic function, an “aperiodic” function is one, which in not periodic. On the other hand, a function is said to be anti-periodic if :

$$f(x+T) = -f(x) \quad \text{for all} \quad x \in D$$

Periodicity and period

In order to determine periodicity and period of a function, we can follow the algorithm as :

1. Put f(x+T) = f(x).
2. If there exists a positive number “T” satisfying equation in “1” and it is independent of “x”, then f(x) is periodic. Otherwise, function, “f(x)” is aperiodic.
3. The least value of “T” is the period of the periodic function.

Example:

Problem : Let f(x) be a function and “k” be a positive real number such that :

$$f(x+k)+f(x)=0 \quad \text{for all} \quad x \in R$$

Prove that f(x) is periodic. Also determine its period.

Solution : The given equation can be re-written as :

$$f(x+k)=-f(x) \quad \text{for all} \quad x \in R$$

Here, our objective is to convert RHS of the equation as f(x). For this, we need to substitute "x" such that RHS function acquires RHS function form. Replacing “x” by “x+k”, we have :

$$\Rightarrow f(x+2k)=-f(x+k) \quad \text{for all} \quad x \in R$$

Combining two equations,

$$\Rightarrow f(x+2k)=-1X-f(x)=f(x) \quad \text{for all} \quad x \in R$$

It means that f(x) is a periodic function and its period is “2k”.

Example:

Problem : Determine period of the function :

$$f(x)=a\sin kx+b\cos kx$$

Solution : The function is sum of two trigonometric functions. We can reduce this function is terms of a single trigonometric function to determine its periodic nature. Let

$$a=r\cos\theta; \quad b=r\sin\theta$$

$$\Rightarrow r=\sqrt{(a^2+b^2)}$$

Substituting in the function, we have :

$$\Rightarrow f(x) = r\cos\theta\sin kx + r\sin\theta\cos kx = r\sin(kx+\theta)$$

This is a periodic function. Also, period of “ag(x)” is same as that of “g(x)”. Therefore, period of “r sin (kx + θ)” is same as that of “sin (kx + θ)”. On the other hand, period of g(ax+b) is equal to the period of g(x), divided by “|a|”. Now, period of “sinx” is “2π”. Hence, period of the given function is :

$$\Rightarrow T = \frac{2\pi}{|k|}$$

Alternatively, we can treat given function as addition of two functions. The period of each term is “2π/|k|”. Applying LCM rule (discussed later), the period of given function is equal to LCM of two periods, which is “2π/|k|”.

Basic periodic functions

Not many of the functions that we encounter are periodic. There are few functions, which are periodic by their very definition. We are, so far, familiar with following periodic functions in this course :

- Constant function, (c)
- Trigonometric functions, (sinx, cosx, tanx etc.)
- Fraction part function, {x}

Six trigonometric functions are most commonly used periodic functions. They are used in various combination to generate other periodic functions. In general, we might not determine periodicity of each function by definition. It is more convenient to know periods of standard functions like that of six trigonometric functions, their integral exponents and certain other standard forms/ functions. Once, we know periods of standard functions, we use different rules, properties and results of periodic functions to determine periods of other functions, which are formed as composition or combination of standard periodic functions.

Constant function

For constant function to be periodic function,

$$f(x+T) = \mathrm{f}(\mathrm{x})$$

By definition of constant function,

$$f(x+T) = \mathrm{f}(\mathrm{x}) = c$$

Clearly, constant function meets the requirement of a periodic function, but there is no definite, fixed or least period. The relation of periodicity, here, holds for any change in x. We, therefore, conclude that constant function is a periodic function without period.

Trigonometric functions

Graphs of trigonometric functions (as described in the module titled trigonometric function) clearly show that periods of sinx, cosx, cosecx and secx are “2π” and that of tanx and cotx are “π”. Here, we shall mathematically determine periods of few of these trigonometric functions, using definition of period.

Sine function

For sinx to be periodic function,

$$\sin(x+T) = (x)$$

$$x+T = n\pi + (-1)^n x; \quad n \in Z$$

The term $(-1)^n$ evaluates to 1 if n is an even integer. In that case,

$$x+T = n\pi + x$$

Clearly, T = nπ, where n is an even integer. The least positive value of “T” i.e. period of the function is :

$$T = 2\pi$$

Cosine function

For cosx to be periodic function,

$$\cos(x + T) = \cos x$$

$$\Rightarrow x + T = 2n\pi \pm x; \quad n \in Z$$

Either,

$$\Rightarrow x + T = 2n\pi + x$$

$$\Rightarrow T = 2n\pi$$

or,

$$\Rightarrow x + T = 2n\pi - x$$

$$\Rightarrow T = 2n\pi - 2x$$

First set of values is independent of “x”. Hence,

$$T = 2n\pi; \quad n \in Z$$

The least positive value of “T” i.e. period of the function is :

$$T = 2\pi$$

Tangent function

For tanx to be periodic function,

$$\tan(x + T) = \tan x$$

$$x + T = n\pi + x; \quad n \in Z$$

Clearly, T = nπ; n∈Z. The least positive value of “T” i.e. period of the function is :

$$T = \pi$$

Fraction part function (FPF)

Fraction part function (FPF) is related to real number "x" and greatest integer function (GIF) as $\{x\} = x - [x]$. We have seen that greatest integer function returns the integer which is either equal to “x” or less than “x”. For understanding the nature of function, let us compute few function values as here :

```
--------------------------------- x [x] x - [x] --
------------------------------- 1 1 0 1.25 1 0.25
1.5 1 0.5 1.75 1 0.75 2 2 0 2.25 2 0.25 2.5 2 0.5
2.75 2 0.75 3 3 0 3.25 3 0.25 3.5 3 0.5 3.75 3
0.75 4 4 0 ---------------------------------
```

From the data in table, we infer that difference as given by “x – [x]” is periodic with a period of “1”. Note that function value repeats for an increment of "1" in the value of "x". We, now, proceed to prove this analytically. Here,

$$f(x) = x - [x]$$

$$\Rightarrow f(x+T) = x + T - [x+T]$$

Let us assume that the given function is indeed a periodic function. Then by definition,

$$f(x+T) = f(x)$$

$$\Rightarrow x + T - [x+T] = x - [x]$$

$$\Rightarrow T = [x+T] - [x]$$

$$\Rightarrow T \in Z$$

Clearly, “T” is an integer as both greatest integer functions return integers. There exists T > 0, which satisfies the equation f(x+T) = f(x). The least positive integer is “1”. Hence, period of the function is “1”.

Arithmetic operations and periodicity

A periodic function can be modified by arithmetic operations on independent variable of the function or function itself. The arithmetic operations involved here are addition, subtraction, multiplication, division and negation. We have studied (read module titled transformation of graphs) these operations and seen that there are different effects on the graph of core function due to these operations. Arithmetic operations on independent variable change input to the function and the graph of core function is transformed horizontally (along x-axis). On the other hand, operations on the function itself change output and the graph of core function is transformed vertically (along. y-axis). The combined input/output arithmetic operations related to function are symbolically represented as :

$$af(bx+c)+d; \quad \text{a,b,c,d} \in Z$$

Important thing to understand here is that periodicity is defined in terms of independent variable, x. A periodic function repeats a set of its values after regular interval of independent variable i.e. x., Clearly, periodicity of a periodic function is not affected by transformations in vertical direction. Hence, arithmetic operations with function involving constants “a” and “d” do not affect periodicity of a periodic function.

Not all arithmetic operations on independent variable will change or affect periodicity. Shifting of core graph due to addition or subtraction results in shifting of the graph as a whole either to the left or right. This operation does not change size and shape of the graph. Thus, addition and subtraction operation involving constant “c” does not affect periodicity of a function. Negation of independent variable, when “b” is negative, results in flipping of the graph without any change in size and shape of the graph. As such, negation of independent variable does not change periodicity either.

It is only the multiplication or division of independent variable x by a positive constant, “b” greater than 1, result in change in size with respect to origin in horizontal direction. The graph shrinks horizontally when independent variable is multiplied by positive constant greater than 1 by the factor which is equal to the multiplier. This means periodicity of graph decreases by the same factor i.e.|b|. The graph stretches horizontally when independent variable is divided by positive constant greater than 1 by the factor which is equal to the divisor. This means periodicity of graph increases by the same factor i.e.|b|. We combine these two observations by saying that period of graph decrease by a factor |b|. Note that magnitude of constant “b” more than 1 represents multiplication and less than represents division.

In the nutshell, if “T” is the period of f(x), then period of function of the form given below id “T/|b|” :

$$af(bx+c)+d; \quad \text{a,b,c,d} \in Z$$

Example:
Problem : What is the period of function :

$$f(x)=3+2\sin\left\{\frac{(\pi x+2)}{3}\right\}$$

Solution : Rearranging, we have :

$$f(x)=3+2\sin\left(\frac{\pi}{3}x+\frac{2}{3}\right)$$

The period of sine function is “2π”. Comparing with function form "$af(bx+c)+d$", magnitude of b i.e. |b| is π/3. Hence, period of the given function is :

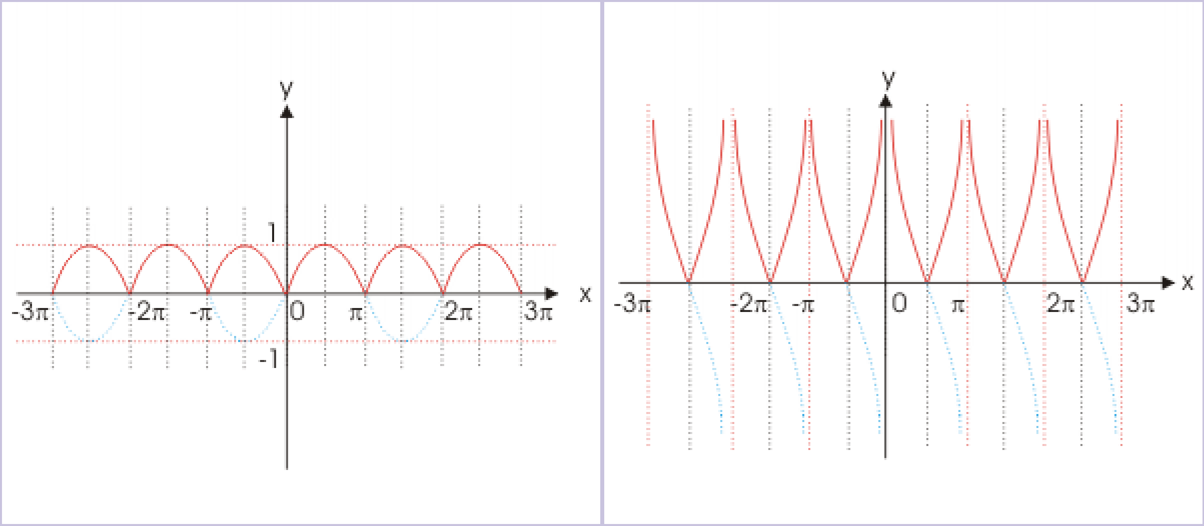
$$\Rightarrow T\prime = \frac{T}{|b|} = \frac{2\pi}{\frac{\pi}{3}} = 6$$

Modulus of trigonometric functions and periods

The graphs of modulus of a function are helpful to determine periods of modulus of trigonometric functions like |sinx|, |cosx|, |tanx| etc. We know that modulus operation on function converts negative function values to positive function values with equal magnitude. As such, we draw graph of modulus function by taking mirror image of the corresponding core graph in x-axis. The graphs of |sinx| and |cotx| are shown here :

Modulus of sine and cotangent functions

Period of function is π. Period of function is π.



From the graphs, we observe that periods of |sinx| and |cotx| are π. Similarly, we find that periods of modulus of all six trigonometric functions are π.

Integral exponentiation of trigonometric function and periods

The periods of trigonometric functions which are raised to integral powers, depend on the nature of exponents. The periods of trigonometric exponentiations are different for even and odd powers. Following results with respect these exponentiated trigonometric functions are useful :

Functions $\sin^n x, \cos^n x, \operatorname{cosec}^n x \quad \text{and} \quad \sec^n x$ are periodic on "R" with period "π" when "n" is even and "2π" when "n" is fraction or odd. On the other hand, Functions $\tan^n x \quad \text{and} \quad \cot^n x$ are periodic on "R" with period "π" whether n is odd or even.

Example:

Problem : Find period of $\sin^2 x$.

Solution : Using trigonometric identity,

$$\Rightarrow \sin^2 x = \frac{1+\cos 2\text{x}}{2}$$

$$\Rightarrow \sin^2 x = \frac{1}{2} + \frac{\cos 2\text{x}}{2}$$

Comparing with $af(bx+c)+d$, the magnitude of "b" i.e. |b| is 2. The period of cosine is 2π. Hence, period of $\sin^2 x$ is :

$$\Rightarrow T = \frac{2\pi}{2} = \pi$$

Example:

Problem : Find period of function :

$$f(x) = \sin^3 x$$

Writing identity for " $\sin^3 x$ ", we have :

$$\Rightarrow f(x) = \sin^3 x = \frac{3\sin x - \sin 3x}{4} = \frac{3}{4}\sin x - \frac{3}{4}\sin 3x$$

We know that period of “ag(x)” is same as that of “g(x)”. The period of first term of “f(x)”, therefore, is equal to the period of “sinx”. Now, period of “sinx” is “2π”. Hence,

$$\Rightarrow T_1 = 2\pi$$

We also know that period of g(ax+b) is equal to the period of g(x), divided by “|a|”. The period of second term of “f(x)”, therefore, is equal to the period of “sinx”, divided by “3”. Now, period of “sinx” is “2π”. Hence,

$$\Rightarrow T_2 = \frac{2\pi}{3}$$

Applying LCM rule,

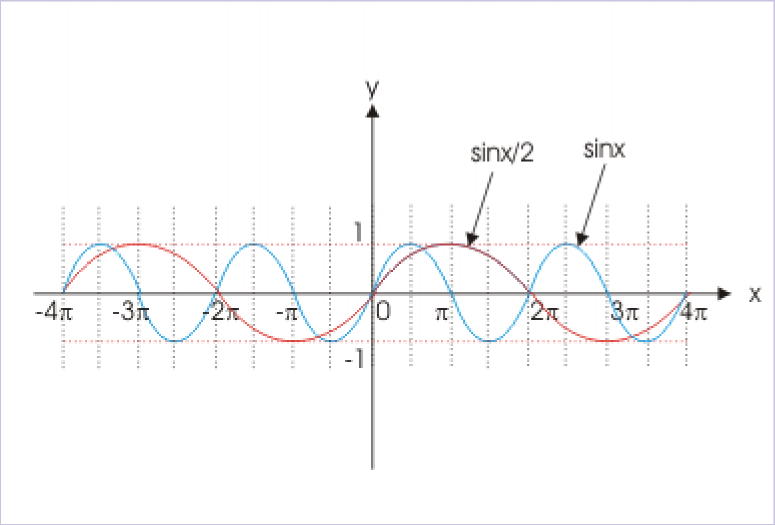
$$\Rightarrow T = \frac{\text{LCM of} \quad 2\pi \quad \text{and} \quad 2\pi}{\text{HCF of} \quad 1 \quad \text{and} \quad 3} = \frac{2\pi}{1} = \pi$$

LCM Rule for periodicity

When two periodic functions are added or subtracted, the resulting function is also a periodic function. The resulting function is periodic when two individual periodic functions being added or subtracted repeat simultaneously. Consider a function,

$$\mathrm{f(x)} = \mathrm{sinx} + \sin\frac{x}{2}$$

The period of sinx is 2π, whereas period of sinx/2 is 4π. The function f(x), therefore, repeats after 4π, which is equal to LCM of (least common multiplier) of the two periods. It is evident from the graph also.
Graphs of sine functions



Two functions repeat values after 4π.

If “a” and “b” are non-zero real number and functions g(x) and h(x) are periodic functions having periods, “ T_1 ” and “ T_2”, then function $f(x) = ag(x) \pm bh(x)$ is also a periodic function. The period of f(x) is LCM of "T_1" and "T_2".

Finding LCM

LCM of integral numbers is obtained easily. There is, however, difficulty in finding LCM when numbers are fractions (like 3/4, 1/3 etc.) or irrational numbers (like π, 2√2 etc.).

For rational fraction, we can find LCM using following formula :

$$\text{LCM} = \frac{\text{LCM of numerators}}{\text{HCF of denominators}}$$

Consider fractions 3/5 and 2/3. The LCM of numerators 3 and 2 is 6. The HCF of denominators is 1. Hence, LCM of two fractions is 6/1 i.e. 6.

This rule also works for irrational numbers of similar type like 2√2/3 , 3√2/5 etc or π/2, 3π/2 etc. However, we can not find LCM of irrational numbers of different kind like 2√2 and π. Similarly, there is no LCM for combination of rational and irrational numbers.

For example, LCM of π/3 and 3π/2 is :

$$\text{LCM} = \frac{\text{LCM of numerators}}{\text{HCF of denominators}} = \frac{\text{LCM of } (\pi, 3\pi)}{\text{HCF of } (2,3)} = \frac{3\pi}{1} = 3\pi$$

Example:
Problem : Find period of

$$f(x) = \sin\left(2\pi x + \frac{\pi}{8}\right) + 2\sin\left(3\pi x + \frac{\pi}{3}\right)$$

Solution : Period of $\sin(2\pi x + \pi/8)$ is :

$$\Rightarrow T_1 = \frac{2\pi}{2\pi} = 1$$

Period of $2\sin(3\pi x + \pi/3)$ is :

$$\Rightarrow T_2 = \frac{2\pi}{3\pi} = \frac{2}{3}$$

LCM of numbers involving fraction is equal to the ratio of LCM of numerators and HCF of denominators. Hence,

$$\text{LCM} = \frac{\text{LCM of numerators}}{\text{HCF of denominators}} = \frac{\text{LCM of } (1,2)}{\text{HCF of } (1,3)} = \frac{2}{1} = 2$$

Exception to LCM rule

LCM rule is not always true. There are exceptions to this rule. We do not apply this rule, when functions are co-functions of each other or when functions are even functions. Further, if individual periods are rational and irrational numbers respectively, then LCM is not defined. As such, this rule can not be applied in such situation as well.

Two functions f(x) and g(y) are cofunctions if x and y are complimentary angles. The functions sinx and consx are cofunctions as :

$$\sin x = \cos\left(\frac{\pi}{2} - x\right)$$

Similarly, |cosx| and |sinx| are cofunctions as :

$$\left|\cos x\right| = \left|\sin\left(\frac{\pi}{2} - x\right)\right|$$

In such cases where LCM rule is not applicable, we proceed to apply definition of periodic function to determine period.

Example:

Problem : Find period of

$$f(x) = |\cos x| + |\sin x|$$

Solution : We know that |cosx| and |sinx| are co-functions. Recall that a function “f” is co-function of a function “g” if f(x) = g(y) where x and y are complementary angles. Hence, we can not apply LCM rule. But, we know that sin(x + π/2) = cosx. This suggests that the function may have the period "π/2". We check this as :

$$f\left(x + \frac{\pi}{2}\right) = \left|\cos\left(x + \frac{\pi}{2}\right)\right| + \left|\sin\left(x + \frac{\pi}{2}\right)\right|$$

$$\Rightarrow f\left(x + \frac{\pi}{2}\right) = \left|-\cos x\right| + \left|\sin x\right| = \left|\cos x\right| + \left|\sin x\right| = f(x)$$

Hence, period is “$\pi/2$”.

It is intuitive here to work with this problem using LCM rule and compare the result. The period of modulus of all six trigonometric functions is π. The periods of |cosx| and |sinx| are π. Now, applying LCM rule, the period of given function is LCM of π and π, which is π.

Example:

Problem : Find period of

$$f(x) = \sin^2 x + \cos^4 x$$

Solution : Here, we see that f(-x) = - f(x).

$$f(\text{-x}) = \sin^2(\text{-x}) + \cos^4(\text{-x}) = \sin^2 x + \cos^4 x = f(x)$$

This means that given function is even function. As such, we can apply LCM rule. We, therefore, proceed to reduce the given function in terms of one trigonometric function type.

$$\Rightarrow f(x) = \sin^2 x + \cos^2 x \left(1 - \sin^2 x\right) = \sin^2 x + \cos^2 x - \sin^2 X \cos^2 x$$

$$\Rightarrow f(x) = 1 - \frac{1}{4}\sin^2 2x = 1 - \frac{1}{4} X \frac{(1 - \cos 4x)}{2}$$

$$\Rightarrow f(x) = 1 - \frac{1}{8} + \frac{(\cos 4x)}{8}$$

$$\Rightarrow T = \frac{2\pi}{4} = \pi/2$$

Note that if we apply LCM rule, then period evaluates to "π".

Important results

The results obtained in earlier sections are summarized here for ready reference.

1: All trigonometric functions are periodic on “R”. The functions $\sin x, \cos x, \sec x \quad \text{and} \quad \operatorname{cosec} x$ have periodicity of "2π". On the other hand, periodicity of $\tan x \quad \text{and} \quad \cot x$ is π.

2: Functions $\sin^n x, \cos^n x, \operatorname{cosec}^n x \quad \text{and} \quad \sec^n x$ are periodic on “R” with period “π” when “n” is even and “2π” when “n” is fraction or odd. On the other hand, Functions $\tan^n x \quad \text{and} \quad \cot^n x$ are periodic on “R” with period “π” whether n is odd or even.

3: Functions $|\sin x|, |\cos x|, |\tan x|, |\operatorname{cosec} x|, |\sec x| \quad \text{and} \quad |\cot x|$ are periodic on “R” with period “π”.

4: A constant function is a periodic function without any fundamental period. For example,

$$f(x) = c$$

$$f(x) = \sin^2 x + \cos^2 x$$

5: If “T” is the period of f(x), then period of function of the form given below is “T/|b|” :

$$af(bx + c) + d; \quad \text{a,b,c,d} \in Z$$

6: If f(x) is a periodic function with period “T” and g(x) is one one function (bijection), then "gof" is also periodic with period “T”.

7: If f(x) is a periodic function with a period T and its domain is a proper subset of domain of g(x), then gof(x) is a periodic function with a period T.

Example:
Problem : Find period of function :

$$f(x) = \sin(x)$$

where {} denotes fraction part function.

Solution : The fraction part function {x} is a periodic function with a period “1”. Its domain is R. On the other hand, sinx is a function having domain R. Therefore, domain of {x} is a proper subset of the domain of sinx. Hence, period of sin{x} is 1.

Example:

Problem : Find period of function $\tan^{-1}\tan x$

Solution : Inverse trigonometric function $\tan^{-1}x$ is one one function in [1,1] and tanx is a periodic function with period π in R. Hence, function $\tan^{-1}\tan x$ function is a periodic function with period π.

Periodicity and periods

Determination of period requires analysis based on properties and methods of functions, including transformation of graphs. We can use following guidelines to determine period :

1 : The definition of periodicity is helpful in determining period of generic function like f(x) under certain condition whose form i.e. rule is not given. The definition is also used to determine whether given function is periodic or not in the first place.

2 : Various generalizations about periods of trigonometric functions, their even and odd exponentiations etc. help to determine periods of individual terms in a trigonometric expression.

3 : Transformation of graphs or functions is an important method to determine periods for function expression, which involves arithmetic operations like addition, subtraction, multiplication, division or negation. These operations are applied either on independent variable or function itself.

4 : One of important concepts for determining period of function involving multiple periodic terms is that individual function should repeat simultaneously. This gives rises to LCM rule. It is defined for terms which appear as sum or difference in the function. However, LCM concept can be extended to division or multiplication of periodic terms as well.

5 : Exceptions to LCM rule are important. Even function and function comprising of cofunctions are two notable exceptions to LCM rule. Besides, LCM of irrational periods of different kinds is not possible. This fact is used to determine periodic nature of function involving irrational periods. If LCM can not be determined, then given function is not periodic in the first place.

6 : Composition of function gof(x) has same period as that of the period of f(x) under certain conditions : either domain of f(x) is proper subset of domain of g(x) or g(x) is a bijection (one-one function).

Periodicity involving generic function

We are required to determine period of a generic function like f(x) based on certain conditional relation. In such situation, we are required to manipulate given condition in such a manner that we ultimately get a relation of type f(x+T)= f(x). Generally, this requires substitution of independent variable with expression which results in existing expressions. This enables us to use given relation repeatedly. This concept is better understood by working with an example.

Example:

Problem : A function f(x) satisfies equation given :

$$f(x+1)+f(x-1)=\sqrt{3}f(x)$$

Determine its period.

Solution : Here, main strategy is to replace "x" such that we get expression on RHS which is same as the expression on LHS. Replacing "x" by "x+1" and replacing "x" by "x-1" separately in given equation, we get two equations :

$$\Rightarrow f(x+2)+f(x)=\sqrt{3}f(x+1)$$

$$\Rightarrow f(x)+f(x-2)=\sqrt{3}f(x-1)$$

Adding these two equations, we get term on RHS, which is same as LHS :

$$\Rightarrow f(x+2)+f(x-2)+2f(x)=\sqrt{3}\left\{f(x+1)+f(x-1)\right\}=3f(x)$$

$$\Rightarrow f(x+2)+f(x-2)=f(x)$$

Replacing "x" by "x+2" in above equation, we have :

$$\Rightarrow f(x+4)+f(x)=f(x+2)$$

Note that we have not replaced “x” by “x-2”, because that yields a related which has argument form of “x-4”. As definition of periodic function involves addition of a positive constant being added to independent variable, we opt to replace “x” by “x+2”. Now, adding two preceding equations,

$$\Rightarrow f(x+4) + f(x-2) = 0$$

Replacing “x” by “x+2” so that one of two term becomes f(x), we have :

$$\Rightarrow f(x+6) + f(x) = 0$$

$$\Rightarrow f(x+6) = -f(x)$$

Replacing “x” by “x+6”, we have :

$$\Rightarrow f(x+12) = -f(x+6) = f(x)$$

$$\Rightarrow T = 12$$

Periodicity involving fraction part function

The period of {x} is 1. We use transformation rule to determine period of various function forms involving arithmetic operations. It means that only coefficient of "x" changes period of FPF.

Example:
Problem : Find period of function :

$$f(x) = 3\sin 2\{3x\} + 5\cos 3\{2x\}$$

where {} denotes fraction part function.
Solution : The period of {x} is 1. Therefore, period of function {3x} is 1/3. Domains of sin2x and FPF functions are R. Thus, domain of FPF is

proper subset of domain of sin2x. Hence, period of 3sin2{3x} is 1/3. On the other hand, period of function {2x} is 1/2. Domains of cos3x and FPF functions are R. Hence, period of 5cos3{2x} is 1/2. LCM of 1/2 and 1/3 is :

$$\Rightarrow LCM = \frac{\text{LCM of} \quad (1,1)}{\text{HCF of} \quad (2,3)} = \frac{1}{1} = 1$$

Periodicity involving radical

Determination of periods involving radicals is evaluated applying LCM rule. We can determine LCM of radicals, if they are of same kind. If radicals involved are of different kinds, then we can not determine LCM. In that case, given function is not periodic.

Example:
Problem : Find period of function :

$$f(x) = 3\sin 2\sqrt{3}x + 2\cos 5\sqrt{3}x$$

Solution : Period of $3\sin 2\sqrt{3}x$ is :

$$\Rightarrow T_1 = \frac{2\pi}{2\sqrt{3}}$$

Period of $2\cos 5\sqrt{3}x$ is :

$$\Rightarrow T_2 = \frac{2\pi}{5\sqrt{3}}$$

Two irrational periods are of same kind. Hence, period of given function is :

$$\Rightarrow LCM = \frac{\text{LCM of} \quad (2\pi, 2\pi)}{\text{HCF of} \quad \left(2\sqrt{3}, 5\sqrt{3}\right)} = \frac{2\pi}{\sqrt{3}}$$

Example:
Problem : Find period of function :

$$f(x) = \sin\sqrt{2}x + \cos\sqrt{3}x$$

Solution : Period of $\sin\sqrt{2}x$ is :

$$\Rightarrow T_1 = \frac{2\pi}{\sqrt{2}}$$

Period of $\cos\sqrt{3}x$ is :

$$\Rightarrow T_2 = \frac{2\pi}{\sqrt{3}}$$

Two irrational periods are not of same kind. Hence, function is not periodic.

Example:
Problem : Determine whether the function given below is periodic?

$$f(x) = \cos\sqrt{x}$$

If the function is periodic, then find its period.
Solution : We need to check periodicity applying definition of periodic function. According to definition of periodic function,

$$f(x + T) = f(x)$$

$$\Rightarrow \cos \sqrt{(x+T)} = \cos \sqrt{x}$$

$$\Rightarrow \sqrt{(x+T)} = 2n\pi \pm \sqrt{x}, \quad n \in Z$$

Squaring both sides and solving for “T”, we have :

$$\Rightarrow (x+T) = \left(2n\pi \pm \sqrt{x}\right)^2$$

$$\Rightarrow T = \left(2n\pi \pm \sqrt{x}\right)^2 - x$$

If we expand the square term, then we find that the expression of “T” is not independent of “x”. Hence, given function is not a periodic function.

Periodicity involving trigonometric combinations

A function comprising of trigonometric functions presents largest possibilities of periodic function. Apart from definition of period, there are many standard results, which can be used to determine periodicity. These results have been discussed in earlier module. In case of function expressions having more than one term, we employ LCM rule to determine period - provided function is not even function or function does not involve cofunctions.

Example:
Problem : Find period of function :

$$f(x) = \sin x + \tan \frac{x}{2}$$

Solution : The function is addition of two trigonometric functions. Each of the functions is periodic. We see that neither sine nor tangent function is even function. Therefore, we apply LCM rule to find the period of the function.

The period of “sin x” is “2π”. Hence,

$$\Rightarrow T_1 = 2\pi$$

We also know that period of g(ax+b) is equal to the period of g(x), divided by “|a|”. The period of second term of “f(x)”, therefore, is equal to the period of “tanx”, divided by “1/2”. Thus, period of “tanx” is “π”.

$$\Rightarrow T_2 = \frac{\pi}{\frac{1}{2}} = 2\pi$$

Now, LCM of “2π” and “2π” is “2π”. Hence,

$$\Rightarrow T = 2\pi$$

Example:
Problem : Determine whether the function given below is periodic?

$$f(x) = x\cos x$$

If the function is periodic, then find its period.
Solution : The function is product of algebraic function “x” and trigonometric function “cosx”. We need to check periodicity applying definition of a periodic function.
Let the function be periodic. Then,

$$f(x+T) = f(x)$$

$$\Rightarrow (x+T)\cos(x+T) = x\cos x$$

$$\Rightarrow T\cos(x+T) = x\cos x - x\cos(x+T) = x\{\cos x - \cos(x+T)\}$$

We see that right hand side is product of algebraic function “x” and trigonometric function. On the left hand side, there is only trigonometric function (apart from “T”, which is a constant). There is no algebraic function in “x” on the left which can cancel "x" on the right. Thus, we conclude that “T” is not independent of “x” and as such give function is not periodic.

Example:

Problem : If the function $f(x) = \sin x + \cos ax$ be periodic, then prove that “a” is rational.

Solution :

Statement of the problem : The function is sum of two trigonometric functions. It is given that the function is a periodic function.

Let “T” be the period, then according to definition :

$$\sin(x+T) + \cos a(x+T) = \sin x + \cos ax$$

Putting, x = 0, we have :

$$\Rightarrow \sin T + \cos aT = 1$$

Putting, x = -T, we have :

$$\Rightarrow \sin 0 + \cos 0 = -\sin T + \cos T$$

$$\Rightarrow -\sin T + \cos aT = 1$$

Subtracting one equation from another,

$$\sin T = 0$$

$$\Rightarrow T = n\pi, \quad n \in Z$$

and two equations :

$$\cos aT = 1$$

$$\Rightarrow aT = 2m\pi, \quad n \in Z$$

Combining two results :

$$\Rightarrow a = \frac{2\mathrm{m}}{n}$$

Hence, “a” is a rational number.

Example:

Problem : Find value of “n”, if period of given function is π/4.

$$f(x) = \frac{4\sin 2nx}{1+\cos^2 nx}; \quad n \in N$$

Solution : The numerator and denominator repeat simultaneously. Hence, period of given function is LCM of the periods of numerator and denominator. Now, period of 4sin2nx is :

$$\Rightarrow T_1 = \frac{2\pi}{2n} = \frac{\pi}{n}$$

Using trigonometric identity, the denominator of the given function is :

$$\Rightarrow 1+\cos{}^2 nx = 1 + \frac{1+\cos 2nx}{2}$$

The period of denominator is :

$$\Rightarrow T_2 = \frac{2\pi}{2n} = \frac{\pi}{n}$$

Hence, period of given function is LCM of π/n and π/n, which is π/n. According to question,

$$\Rightarrow \frac{\pi}{n} = \frac{\pi}{4}$$

$$\Rightarrow n = 4$$

Example:

Problem : Find period of function :

$$f(x) = \sin^3 x + \sin^2\left(x + \frac{\pi}{3}\right) - \cos x \cos\left(x + \frac{\pi}{3}\right)$$

Solution : The given function is not an even function as $\sin^3(-x) = -\sin^3 x$. We can use LCM rule to determine period of given function. The periods of $\sin^3 x$ and $\sin^2\left(x + \frac{\pi}{3}\right)$ are 2π and π respectively. Using trigonometric identity,

$$\cos x \cos\left(x + \frac{\pi}{3}\right) = \frac{1}{2}\cos\left(2x + \frac{\pi}{3}\right) + \cos\left(\frac{\pi}{3}\right)$$

$$\Rightarrow \cos x \cos\left(x + \frac{\pi}{3}\right) = \frac{1}{2}\cos\left(2x + \frac{\pi}{3}\right) + \frac{1}{2}$$

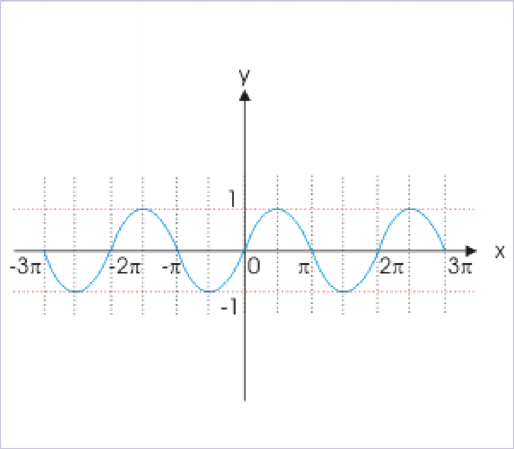
The period of cos (2x+π/3) is 2π/2 = π. Now, periods of three terms of given function are 2π,π and π. Hence, period of given function is LCM of three terms i.e. 2π.

Monotonic functions

The term “monotonic” conveys the meaning of maintaining order or the sense of “no change”. In the context of function, we think a monotonic function as the one whose successive values are increasing, decreasing or constant. There is a sense of maintaining order of function values as independent variable changes. These aspects are pictorially evident on the graph of a function. In a general case, a function may or may not maintain its order of change in its domain i.e. in the overall context. However, we can always identify monotonic behavior in an appropriately chosen subset of domain – unless it is a point function or a singleton.

Consider the graph of sine function. As a whole, the function is not monotonic as the order of the function is not preserved over the domain of the function, which is “R”. However, if we consider an interval, say, between “0” and “π/2”, then we find that function keeps increasing with the increasing independent variable. Therefore, sine function is monotonic in this interval.

Monotonic function



The sine function is monotonic in certain intervals.

On the other hand, a linear polynomial function represents a straight line, which maintains its monotonic nature through out its domain. The monotonic nature of a function, therefore, is investigated in a suitable interval, which is either domain or its subset. We shall refer this interval as X to illustrate the concept in this module. From the point of view of monotonic behavior, we classify function in following categories :

1: Constant function : Function values does not change as independent variable varies.

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) = f(x_2), \quad \text{for all} \quad x_1, x_2 \in X.$$

2: Strictly increasing: Function value change as independent variable varies in accordance with following condition :

$$\text{If}\quad x_1 < x_2 \quad\text{then}\quad f(x_1) < f(x_2),\quad \text{for all}\quad x_1, x_2 \in X.$$

3: Non-decreasing or increasing : Function value change as independent variable varies in accordance with following condition :

$$\text{If}\quad x_1 < x_2 \quad\text{then}\quad f(x_1) \le f(x_2),\quad \text{for all}\quad x_1, x_2 \in X.$$

4: Strictly decreasing: Function value change as independent variable varies in accordance with following condition :

$$\text{If}\quad x_1 < x_2 \quad\text{then}\quad f(x_1) > f(x_2),\quad \text{for all}\quad x_1, x_2 \in X.$$

5: Non-increasing or decreasing : Function value changes as independent variable varies in accordance with following condition :

$$\text{If}\quad x_1 < x_2 \quad\text{then}\quad f(x_1) \ge f(x_2),\quad \text{for all}\quad x_1, x_2 \in X.$$

There is one ambiguity in the definition of classification presented above. According to the definition, a constant function is an increasing, decreasing or both kinds of function. Clearly, this interpretation is wrong and is an exception. An increasing or non-decreasing class actually captures the notion of an overall increasing function, which is intermittently constant and thereby distinguishes this class from strictly increasing order. Similarly, a decreasing or non-increasing class actually captures the notion of an overall decreasing function, which is intermittently constant and thereby distinguishes this class from strictly decreasing order.

Note : It may confound clarity, but we should know that there is another classification. In this classification (i) "strictly increasing" is known simply as "increasing", (ii) "strictly decreasing" is known simply as "decreasing", (iii) "increasing" is known as "monotonically increasing" and (iv) "decreasing" is known as "monotonically decreasing". Clearly, this classification is not the same as what is given here. The best way to deal with this situation is to ignore this confusion and be explicit in what we mean. Saying "strictly increasing" for example ensures that equality of function values is not allowed. Similarly, saying "non-decreasing" ensures that function values do not decrease. We shall try to adhere to this explicit classification to the extent possible.

Derivative and nature of function

We shall learn subsequently that first derivative of a function is defined in terms of ratio of the differences between two successive values of function and that of independent variable, however small is the difference in independent variable. It is easy to visualize, then, that if we know the nature of derivative in a given interval, then we can determine monotonic behavior of the function as well.

$$f\prime(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

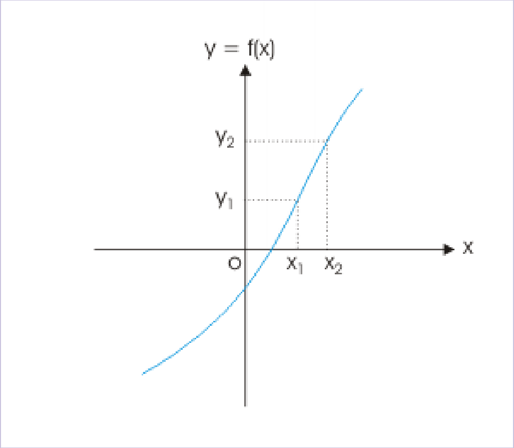
Depending on the monotonic nature of function, the relative values of f(h) and f(x+h) are different and so the sign of first derivative.

Strictly increasing function

The successive value of function increases as the value of the independent variable increases. In other words, the preceding values are less than successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_{2,} \quad \text{then} \quad f(x_1) < f(x_2)$$

Increasing function



The successive function value is greater than previous value.

Example:
Problem : Determine monotonic nature of the function in the interval [0,∞).

$$y = x^2$$

Solution : Let x_1 and x_2 belong to the interval [0,∞) such that $x_1 < x_2$. Multiplying inequality with x_1 (a positive number) does not change the nature of inequality :

$$\Rightarrow x_1^2 < x_1 x_2$$

Multiplying inequality with x_2 (a positive number) does not change the nature of inequality :

$$\Rightarrow x_1 x_2 < x_2^2$$

Combining two inequalities,

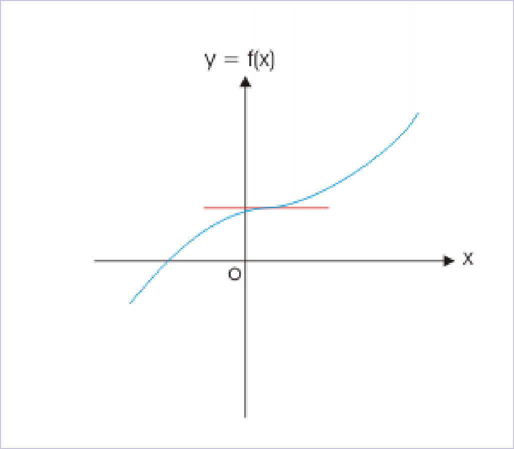
$$\Rightarrow x_1^2 < x_2^2$$

$$\Rightarrow f(x_1) < f(x_2)$$

Thus, given function is strictly decreasing in [0,∞).

As f(x_1)<f(x_2) for all x_1,x_2∈X, the difference "f(x+h) – f(x)" is positive for "h", however small. This implies that the first derivative of function is positive. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for couple of x values, while curve is continuously increasing in the interval. It means that first derivative can be equal to zero for few points in the interval in which it is strictly increasing. This is clear from the figure given here,

Strictly increasing function



The first derivative can be zero at certain point(s) - but not in continuous interval.

Thus, for strictly increasing function,

$$f\prime(x) \geq 0; \quad \text{Equality sign holds for points only - not on a continuous section in X}$$

For strictly increasing function, if x_1<x_2, then f(x_1) < f(x_2), for all x_1,x_2∈X. It means that all distinct x values correspond to distinct y values and vice-versa. Therefore, strictly increasing function is one-one function i.e. a bijection and hence "invertible". In other words, if a function has strict increasing order, then it is invertible. Mathematically, we say that if f'(x) ≥0; (equality holding for points only), x∈X, then function is invertible in X. For example, consider sine function,

$$f(x) = \sin x$$

$$\Rightarrow f\prime(x) = \cos x$$

We know that cosx is positive in the interval [-π/2,π/2]. Hence sine function is a strictly increasing function in [-π/2,π/2] and is invertible. Recall that inverse sine function is defined in this interval.

The order of a function provides an easy technique to determine range of a continuous function, corresponding to a given domain interval. For example, if domain of a continuously increasing function, f(x), is [x_1, x_2], then the least value of the function is f(x_1) and greatest value of the function is f(x_2). Hence, range of the function is :

$$\text{Range} = [f(x_1), f(x_2)]$$

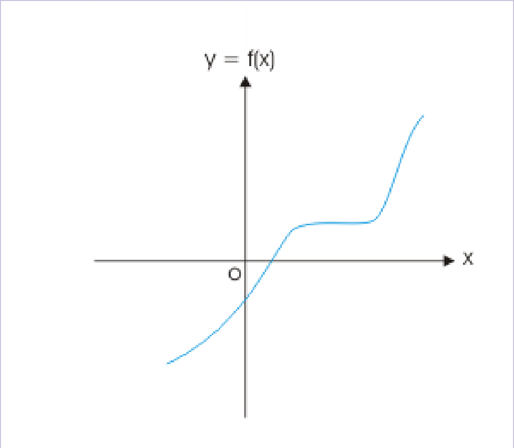
We shall study this aspect of finding range in detail in a separate module.

Non-decreasing function or increasing

The successive value of function increases or remains constant as the value of the independent variable increases. In other words, the preceding values are less than or equal to successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) \leq f(x_2)$$

Non decreasing function



The successive function value is greater than or equal to previous value.

As f(x_1)≤f(x_2) for all x_1,x_2∈X, the difference "f(x+h) – f(x)" is non-negative for "h", however small. This implies that the first derivative of function is non-negative. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for a subset of X, while curve is increasing overall in the interval. It means that first derivative can be equal to zero points or sub-intervals in which it is increasing. Thus, for non-decreasing function,

$$f\prime(x) \geq 0; \quad \text{Equality sign holds for few points or a continuous section in X}$$

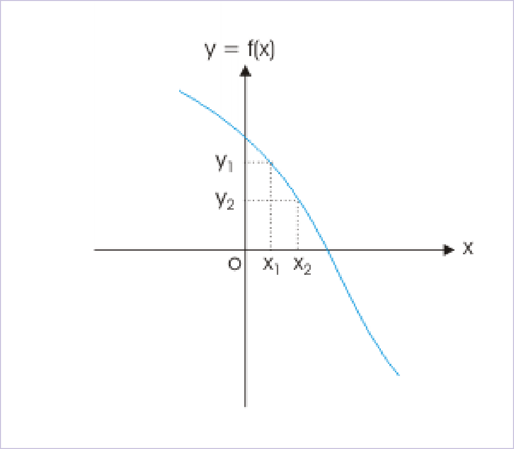
For increasing function, if x_1<x_2, then f(x_1) ≤ f(x_2), for all x_1,x_2∈X. This means that there may be same function values for different values of x. This is "many one" relation and as such function is not invertible in X.

Strictly decreasing function

The successive value of function decreases as the value of the independent variable increases. In other words, the preceding values are greater than successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) > f(x_2)$$

Strictly decreasing function



The successive function value is less than previous value.

Example:

Problem : Determine monotonic nature of the function in the interval (-∞,0].

$$y = x^2$$

Solution : Let x_1 and x_2 belong to the interval [0,∞) such that x_1 < x_2. Multiplying inequality with x_1 (a negative number) changes the nature of inequality :

$$\Rightarrow x_1^2 > x_1 x_2$$

Multiplying inequality with x_2 (a negative number) changes the nature of inequality :

$$\Rightarrow x_1 x_2 > x_2^2$$

Combining two inequalities,

$$\Rightarrow x_1^2 > x_2^2$$

$$\Rightarrow f(x_1) > f(x_2)$$

Thus, given function is strictly decreasing in (-∞,0].

As f(x_1)>f(x_2) for all x_1,x_2∈X, the difference “f(x+h) – f(x)” is negative for “h”, however small. This implies that the first derivative of function is negative. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for couple of x values, while curve is continuously decreasing in the interval. It means that first derivative can be equal to zero for few points in the interval in which it is strictly decreasing. Thus, for strictly decreasing function,

$$f\prime(x) \le 0; \quad \text{Equality sign holds for points only - not on a continuous section in X}$$

For strictly decreasing function, if x_1<x_2, then f(x_1) > f(x_2), for all x_1,x_2∈X. It means that all distinct x values correspond to distinct y values and vice-versa. Therefore, strictly decreasing function is one-one function i.e. a bijection and hence “invertible”. In other words, if a function has strict decreasing order, then it is invertible. Mathematically, we say that if f’(x) ≤ 0 (equality holding for points only); x∈X, then function is invertible in X. For example, consider sine function,

$$f(x) = \sin x$$

$$\Rightarrow f\prime(x) = \cos x$$

We know that cosx is negative in the interval [π/2, 3π/2]. Hence sine function is a strictly decreasing function in [π/2, 3π/2] and is invertible. Recall though that inverse sine function is not defined in this interval, but in basic interval about origin [-π/2,π/2].

The order of a function provides an easy technique to determine range of a continuous function, corresponding to a given domain interval. For example, if domain of a continuously decreasing function, f(x), is [x_1, x_2], then the least value of the function is f(x_2) and greatest value of the function is f(x_1). Hence, range of the function is :

$$\text{Range} = [f(x_2), f(x_1)]$$

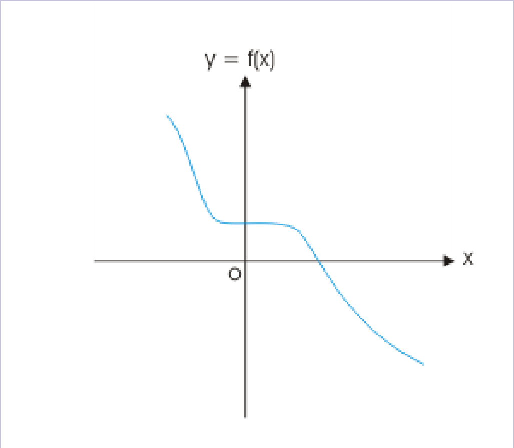
We shall study this aspect of finding range in detail in a separate module.

Non-increasing function or decreasing

The successive value of function decreases or remains constant as the value of the independent variable increases. In other words, the preceding values are greater than or equal to successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) \ge f(x_2)$$

Non-increasing function or decreasing



The successive function value is less than or equal to previous value.

As f(x_1)≥ f(x_2) for all x_1,x_2∈X, the difference “f(x+h) – f(x)” is non-positive for “h”, however small. This implies that the first derivative of function is non-positive. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for a subset of X, while curve is decreasing overall in the interval. It means that first derivative can be equal to zero at points or in sub-intervals in which it is decreasing. Thus, for non-decreasing function,

$$f\prime(x) \le 0; \quad \text{Equality sign holds for few points or a continuous section in X}$$

For decreasing function, if x_1<x_2, then f(x_1) ≥ f(x_2), for all x_1,x_2∈X. This means that there may be same function values for different values of x. This is “many one” relation and as such function is not invertible in X.

Increasing and decreasing intervals

A function is strictly increasing, strictly decreasing, non-decreasing and non-increasing in a suitably selected interval in the domain of the function. We have seen that a linear algebraic function maintains order of change throughout its domain. The order of change, however, may not be maintained for higher degree algebraic and other functions in its domain. We shall, therefore, determine monotonic nature in sub-intervals or domain as the case be.

One of the fundamental ways to determine nature of function is by comparing function values corresponding to two independent values (x_1 and x_2). This technique to determine nature of function works for linear and some simple function forms and is not useful for functions more complex in nature. In this module, we shall develop an algorithm based on derivative of function for determining nature of function in different intevals.

In the discussion about monotonic function in earlier module, we observed that order of change in function values is related to sign of the derivative of function. The task of finding increasing and decreasing intervals is, therefore, about finding sign of derivative of function in different intervals and determining points or intervals where derivative turns zero.

Nature of function and intervals

The steps for determining intervals are given as under :

1: Determine derivative of given function i.e. f’(x).

2: Determine sign of derivative in different intervals.

3: Determine monotonic nature of function in accordance with following categorization :

$$f'(x) \geq 0 : \quad \text{equality holding for points only – strictly increasing interval}$$

$$f'(x) \geq 0 : \quad \text{equality holding for subsections also – non-decreasing or increasing interval}$$

$$f'(x) \leq 0 : \quad \text{equality holding for points only – strictly decreasing interval}$$

$$f'(x) \leq 0 : \quad \text{equality holding for subsections also – non-increasing or decreasing interval}$$

5: The interval is open “( )” at end points, if function is not continuous at end points. However, interval is close “[]” at end points, if function is continuous at end points.

In order to illustrate the steps, we consider a function,

$$f(x) = x^2 - x$$

Its first derivative is :

$$f'(x) = 2x - 1$$

Here, critical point is 1/2. First derivative, f’(x), is positive for x>1/2 and negative for x<1/2. The signs of derivative are strict inequalities. It means that function is either strictly increasing or strictly decreasing in the open intervals. We know that infinity end is an open end. But, function is continuous in the given interval. Hence, we can include end point x=1/2. Further, since derivative is zero at x=1/2 i.e. at a single point, function remains strictly increasing or decreasing.

$$\text{Strictly increasing interval} = \left[-\infty, \frac{1}{2}\right]$$

$$\text{Strictly decreasing interval} = \left[\frac{1}{2}, \infty\right]$$

Algebraic functions

Derivative of algebraic function is also algebraic. In order to determine sign of derivative, we use sign scheme or wavy curve method, wherever expressions in derivative can be factorized.

Example:
Problem : Determine monotonic nature of function in different intervals :

$$f(x) = 3x^4 - x^3$$

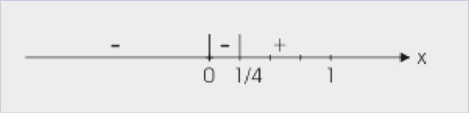
Solution : Its first derivative is :

$$\Rightarrow f'(x) = 12x^3 - 3x^2 = 3x^2(4x - 1)$$

Here, critical points are 0,0,1/4. We have taken 0 twice as we need to write given function in terms of factors as :

$$\Rightarrow f'(x) = 3(x - 0)(x - 0)(4x - 1)$$

Since zero is repeated even times, derivative does not change at x=0. The sign scheme is shown in the figure. First derivative, f’(x), is positive for x>1/4 and negative for x<1/4. Derivative is zero at x=0 and 1/4 i.e. at points only. Clearly, the monotonic nature is "strict" in these intervals. But, function is continuous in the given interval. Hence, we include end point also :
Sign diagram



Increasing and decreasing intervals.

$$\text{Strictly increasing interval} = \left[\frac{1}{4}, \infty\right)$$

$$\text{Strictly decreasing interval} = \left(-\infty, \frac{1}{4}\right]$$

Example:

Problem : Determine monotonic nature of function :

$$f(x) = \sqrt{\left(\frac{\pi^2}{4} - x^2\right)}$$

Solution : The domain of derivative, f(x), is not R. We first need to find its domain and then determine sign of its derivative within the domain. The expression in the square root in denominator is non-negative. Hence,

$$\Rightarrow \frac{\pi^2}{4} - x^2 \geq 0$$

$$\Rightarrow x \in \left[-\frac{\pi}{2,} \frac{\pi}{2}\right]$$

Its first derivative is :

$$\Rightarrow f\prime(x) = -\frac{2x}{2\sqrt{\left(\frac{\pi^2}{4} - x^2\right)}} = -\frac{x}{\sqrt{\left(\frac{\pi^2}{4} - x^2\right)}}$$

We know that square root is a positive number. It means that sign of derivative will solely depend on the sign of “x” in the numerator. Clearly, derivative is positive for x<0 and negative for x>0. Derivative is zero at x= 0 i.e. at a single point only. But, function is continuous in the given interval. Hence, we include end points as well :

$$\text{Strictly increasing interval} = \left[-\frac{\pi}{2}, 0\right]$$

$$\text{Strictly decreasing interval} = \left[0, \frac{\pi}{2}\right]$$

Example:

Problem : Determine monotonic nature of function :

$$f(x) = 2x^3 + 3x^2 - 12x + 1$$

Solution : The first derivative of the given polynomial function is :

$$\Rightarrow f'(x) = 2X3x^2 + 3X2x - 12 = 6x^2 + 6x - 12$$

Clearly, the derivative is a quadratic function. We can determine the sign of the quadratic expression, using sign scheme for quadratic expression. Now, the roots of the corresponding quadratic equation when equated to zero is obtained as :

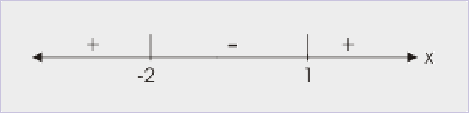
$$\Rightarrow 6x^2 + 6x - 12 = 0 \quad \Rightarrow x^2 + x - 2 = 0 \quad \Rightarrow x^2 + 2x - x - 2 = 0$$

$$\Rightarrow x(x+2) - 1(x+2) = 0 \quad \Rightarrow (x-1)(x+2) = 0$$

$$\Rightarrow x = 1, -2$$

Here, coefficient of “x^2” is positive. Hence, sign of the middle interval is negative and side intervals are positive.

Intervals



Increasing and decreasing intervals

Since cubic polynomial is a continuous function, we can include end points also in the interval :

$$\text{Strictly increasing interval} = (-\infty, -2] \quad \cup \quad [1,\infty)$$

$$\text{Strictly decreasing interval} = [-2,1]$$

Trigonometric function

Derivative of trigonometric function is also trigonometric function. We can determine nature of derivative in two ways. We use trigonometric values to determine nature of sign of first derivative. For this, we make use of sign and value diagram. Alternatively, we find zeroes of trigonometric function and knowing that some of these functions (sine, cosine etc.) changes sign across x-axis and are continuous functions, we can find sign of derivative as required.

Example:

Problem : Determine sub-intervals of [0,π/2] in which given function is (i) strictly increasing and (ii) strictly decreasing.

$$f(x) = \cos 3x$$

Solution : Its first derivative is :

$$f\prime(x) = -3\sin 3x$$

Corresponding to given interval [0,π/2], argument to sine function is [0,3π/2]. Sine function is positive in first two quadrants [0, π] and negative in third quadrant [π,3π/2].
Corresponding to these argument values, sin3x is positive in [0,π/3] and negative in [π/3, π/2] of the given interval. But, negative sign precedes 3sin3x. Hence, derivative is negative in [0,π/3] and positive in [π/3, π/2] of the given interval.
Alternatively, we can find zero of sin3x as :

$$-3\sin 3x = 0$$

$$\Rightarrow \sin 3x = 0$$

$$\Rightarrow 3x = n\pi; \quad n \in Z$$

$$x = \frac{n\pi}{3}; \quad n \in Z$$

Thus, there is one zero at x=π/3 for n=1, in the interval (0,π/2). To test sign, we put x = π/4 in -3sin3x, we have -3sin 3π/4 < 0. Hence, derivative is negative in [0,π/3] and positive in [π/3, π/2]. Further since sine function is a continuous function, we include end points :

$$\text{Strictly decreasing interval} = \left[0, \frac{\pi}{3}\right]$$

$$\text{Strictly increasing interval} = \left[\frac{\pi}{3,}\frac{\pi}{2}\right]$$

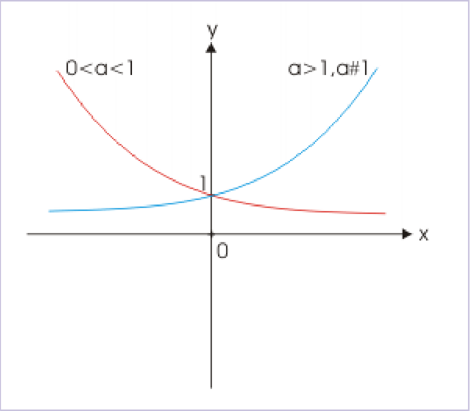
Logarithmic and exponential functions

Logarithmic and exponential functions are continuous in their domain intervals. Important point is that the derivative of exponential function is also an exponential function and it is always positive for base a>0 and a≠1. This fact is also evident from its graphs. On the other hand, derivative of logarithmic function depends on the nature of its argument. The basic derivatives are :

$$f(x) = e^x \quad \Rightarrow f\prime(x) = e^x$$

$$f(x) = \log_e x \quad \Rightarrow f\prime(x) = \frac{1}{x}$$

Graphs of exponential functions



Graphs of exponential functions
for different bases

Example:
Problem : Determine monotonic nature of function :

$$f(x) = (1 + x)e^x$$

Solution : Its first derivative is :

$$f\prime(x) = e^x + (1 + x)e^x = (x + 2)e^x$$

Since e^x is positive for all values of x, the derivative is zero if :

$$(x + 2) = 0$$

$$\Rightarrow x = -2$$

For a test point, $x = -3, \quad f\prime(x) = -1e^{-3} < 0$. Further, function is continuous in R. Hence,

$$\text{Strictly increasing interval} = [-2,\infty)$$

$$\text{Strictly decreasing interval} = (-\infty, -2]$$

Exercises

Exercise:

Problem: Determine monotonic nature of function :

$$f(x) = -2x^3 - 9x^2 + 12x + 7$$

Solution:

Hint : Critical points are 1 and 2.

$$\text{Strictly increasing interval} = [1{,}2]$$

$$\text{Strictly decreasing interval} = (-\infty, 1] \cup [2{,}\infty)$$

Exercise:

Problem:

Determine sub-intervals of [0,π/2] in which given function is (i) strictly increasing and (ii) strictly decreasing.

$$f(x) = \sin 3x$$

Solution:

Its first derivative is :

$$\Rightarrow f\prime(x) = 3\cos 3x$$

Corresponding to given interval [0,π/2], argument to cosine function is [0,3π/2]. Cosine function is positive in first quadrants [0, π/2] and negative in second and third quadrant [π/2,3π/2]. Corresponding to these argument values, 3cos3x is positive in [0,π/6] and negative in [π/6, π/2] of the given interval. Hence, derivative is positive in [0,π/6] and negative in [π/6, π/2].

Alternatively, we can find zero of 3cos3x as :

$$\cos 3x = 0$$

$$\Rightarrow 3x = (2n+1)\frac{\pi}{2}; \quad n \in Z$$

$$\Rightarrow x = (2n+1)\frac{\pi}{6}; \quad n \in Z$$

Thus, there is one zero, x=π/6, in the given interval [0,π/2]. To test sign, we put x = π/4 in 3cos3x, we have 3cos 3π/4 < 0. Hence, derivative is positive in [0,π/6] and negative in [π/6, π/2].

$$\text{Strictly increasing interval} = \left[0, \frac{\pi}{6}\right]$$

$$\text{Strictly decreasing interval} = \left[\frac{\pi}{6}, \frac{\pi}{2}\right]$$

Exercise:

Problem: Determine monotonic nature of function :

$$f(x) = (x-1)e^{x} + 1$$

Solution:

Its first derivative is :

$$\Rightarrow f'(x) = e^{x} + (x-1)e^{x} = e^{x}$$

Since e^{x} is positive for all values of x, the derivative is zero for all x. Hence, given function is strictly increasing in the interval of real number R.

Minimum and maximum values

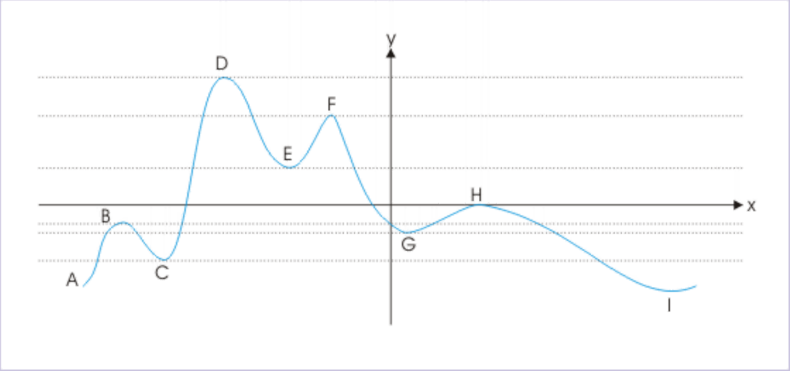
In general context, a function may have multiple minimum and maximum values in the domain of function. These minimum and maximum values are “local” minimum and maximum, which belongs to finite sub-intervals within the domain of function. The least minimum and greatest maximum in the domain of function are “global” minimum and maximum respectively in the entire domain of the function. Clearly, least and greatest values are one of the local minimum and maximum values. The minimum and maximum, which are not global, are also known as “relative” minimum and maximum.

Note : This module contains certain concepts relating to continuity, limits and differentiation, which we have not covered in this course. The topic is dealt here because minimum, maximum, least, greatest and range are important attributes of a function and its study is required to complete the discussion on function.

Important observations and definitions

Let us consider a very general graphic representation of a function. Following observations can easily be made by observing the graph :

Graph of a function



There are multiple minimum and maximum values.

1: A function may have local minimum (C, E, G, I) and maximum (B,D,F,H) at more than one point.

2: It is not possible to determine global minimum and maximum unless we know function values corresponding to all values of x in the domain of function. Note that graph above can be defined to any value beyond A.

3: Local minimum at a point (E) can be greater than local maximum at other points (B and H).

4: If function is continuous in an interval, then pair of minimum and maximum in any order occur alternatively (B,C), (C,D), (D,E), (E,F) , (F,G) , (G,H) , (H,I).

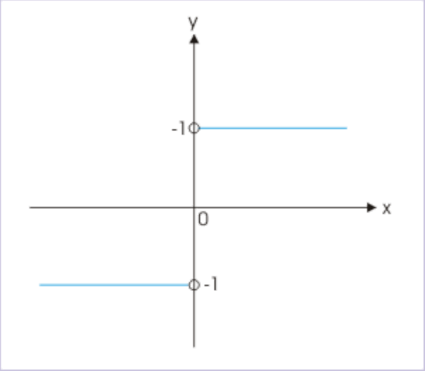
5: A function can not have minimum and maximum at points where function is not defined. Consider a rational function, which is not defined at x=1.

$$f(x) = \frac{1}{x-1}; \quad x \neq 1$$

Similarly, a function below is not defined at x=0.

`| x=1; x>0 f(x) = | |x = -1; x<0`

Graph of function



Function is not at x=0.

Minimum and maximum of function can not occur at points where function is not defined, because there is no function value corresponding to undefined points. We should understand that undefined points or intervals are not part of domain - thus not part of function definition. On the other hand, minimum and maximum are consideration within the domain of function and as such undefined points or intervals should not be considered in the first place. Non-occurance of minimum and maximum in this context, however, has been included here to emphasize this fact.

6: A function can have minimum and maximum at points where it is discontinuous. Consider fraction part function in the finite domain. The function is not continuous at x=1, but minimum occurs at this point (recall its graph).

7: A function can have minimum and maximum at points where it is continuous but not differentiable. In other words, maximum and minimum can occur at corners. For example, modulus function |x| has its only minimum at corner point at x=0 (recall its graph).

Extreme value or extremum

Extreme value or extremum is either a minimum or maximum value. A function, f(x), has a extremum at x=e, if it has either a minimum or maximum value at that point.

Critical points

Critical points are those points where minimum or maximum of a function can occur. We see that minimum and maximum of function can occur at following points :

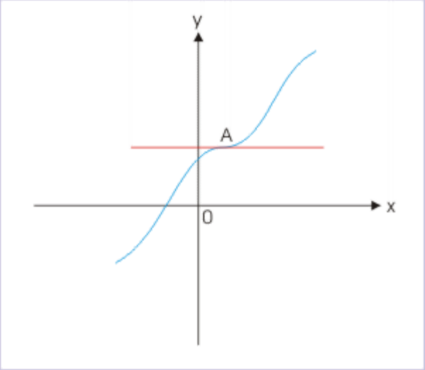
(a) Points on the graph of function, where derivative of function is zero. At these points, function is continuous, limit of function exists and tangent to the curve is parallel to x-axis.

(b) Points where function is continuous but not differentiable. Limit of function exits at those points and are equal to function values. Consider, for example, the corner of modulus function graph at x=0. Minimum of function exist at the corner point i.e at x=0.

(c) Points where function is discontinuous (note that discontinuous is not undefined). A function has function value at the point where it is discontinuous. Neither limit nor derivative exists at discontinuities. Example : piece-wise defined functions like greatest integer function, fraction part function etc.

We can summarize that critical points are those points where (i) derivative of function does not exist or (ii) derivative of function is equal to zero. The first statement covers the cases described at (b) and (c) above. The second statement covers the case described at (a). We should, however, be careful to interpret definition of critical points. These are points where minimum and maximum "can" exist – not that they will exist. Consider the graph shown below, which has an inflexion point at "A". The tangent crosses through the graph at inflexion point. In the illustration, tangent is also parallel to x-axis. The derivative of function, therefore, is zero. But "A" is neither a minimum nor a maximum.

Graph of function



"A" is neither a minimum nor a maximum.

Thus, minimum or maximum of function occur necessarily at critical points, but not all critical points correspond to minimum or maximum.

Note : We need to underline that concept of critical points as explained above is different to the concept of critical points used in drawing sign scheme/ diagram.

Graphical view

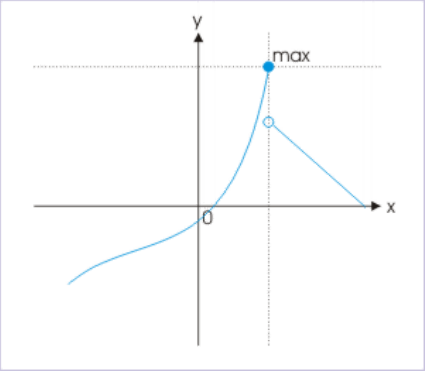
There are mathematical frameworks to describe and understand nature of function with respect to minimum and maximum. We can, however, consider a graphical but effective description that may help us understand occurrence of minimum and maximum values. We need to understand one simple fact that we can have graphs of any nature except for two situations :

1: function is not defined at certain points or in sub-intervals.

2: function can not be one-many relation. In this case, the given relation is not a function in the first place.

Clearly, there exists possibility of minimum and maximum at all points on the continuous portion of function where derivative is zero and at points where curve is discontinuous. This gives us a pictorially way to visualize where minimum and maximum can occur. The figure, here, shows one such maximum value at dicontinuity.

Graph of function



Maximum value at a discontinuity.

Relative or local minimum and maximum

The idea of local or relative minimum and maximum is clearly understood from graphical representation. The minimum function value at a point is least in the immediate neighborhood where minimum occurs. A function has a relative minimum at a point x=m, if function values in the immediate neighborhood on either side of point are less than the value at the point. To be precise, the immediate neighborhood needs to be infinitesimally close. Mathematically,

$$f(m) < f(m+h) \quad \text{and} \quad f(m) < f(m-h) \quad \text{as} \quad h \rightarrow 0$$

The maximum function value at a point is greatest in the immediate neighborhood where maximum occurs. A function has a relative maximum at a point x=m, if function values in the immediate neighborhood on either side of point are greater than the value at the point. To be precise, the immediate neighborhood needs to be infinitesimally close. Mathematically,

$$f(m) > f(m+h) \quad \text{and} \quad f(m) > f(m-h) \quad \text{as} \quad h \rightarrow 0$$

Global minimum and maximum

Global minimum is also known by “least value” or “absolute minimum”. A function has one global minimum in the domain [a,b]. Global minimum, f(l), is either less than or equal to all function values in the domain. Thus,

$$f(l) \le f(x) \quad \text{for all} \quad x \in [a,b]$$

If the domain interval is open like (a,b), then global minimum, f(l), also needs to be less than or equal to function value, which is infinitesimally close to boundary values. It is because open interval by virtue of its inequality does not ensure this. What we mean that it does not indicate how close “x” is to the boundary values. Hence,

$$f(l) \le f(x) \quad \text{for all} \quad x \in (a,b)$$

$$f(l) \le \lim_{x \rightarrow a+0} f(x)$$

$$f(l) \le \lim_{x \to b-0} f(x)$$

Similarly, global maximum is also known by “greatest value” and “absolute maximum”. A function has one global maximum in the domain [a,b]. Global maximum, f(g), is either greater than or equal to all function values in the domain. Thus,

$$f(g) \ge f(x) \quad \text{for all} \quad x \in [a, b]$$

If the domain interval is open like (a,b), then global maximum, f(m), also needs to be greater than or equal to function value, which is infinitesimally close to boundary values. It is because open interval by virtue of its inequality does not ensure this. Hence,

$$f(g) \ge f(x) \quad \text{for all} \quad x \in (a, b)$$

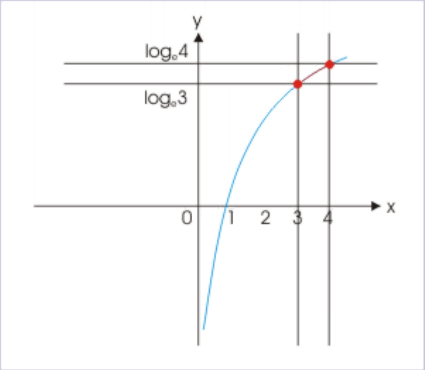
$$f(g) \ge \lim_{x \to a+0} f(x)$$

$$f(g) \ge \lim_{x \to b-0} f(x)$$

Domain interval

Nature of domain interval plays an important role in deciding about occurrence of minimum and maximum and their nature. In order to understand this, we need to first understand that the notion of very large positive value and concept of maximum are two different concepts. Similarly, the notion of very large negative value and concept of minimum are two different concepts. The main difference is that very large negative or positive values are not finite but extremums are finite. Consider a natural logarithmic graph of $\log_e x$. It extends from negative infinity to positive infinity, if base is greater than 1. The function is a strictly increasing function in its entire domain. As such, it has not a single minimum or maximum. The extremely large values at the domain ends can not be considered to be extremum as we can always have function values greater or less than one considered to be maximum or minimum. This argument is valid for behavior of functions near end points of an open interval domain. There can always be values greater or smaller than one considered.

Definite sub-interval of logarithmic function



Definite sub-interval of logarithmic function

However, nature of graph with respect to extremum immediately changes when we define same logarithmic function in a closed interval say [3,4], then $\log_e 3$ and $\log_e 4$ are the respective local minimum and maximum. Incidentally since function is strictly increasing in the domain and hence in the sub-interval, these extremums are global i.e. end values of function are global minimum and maximum in the new domain of the function.

Above argument is valid for all continuous function which may have varying combination of increasing and decreasing trends within the domain of function. The function values at end points of a closed interval are extremums (minimum or maximum) - may not be least or greatest. In the general case, there may be more minimum and maximum values apart from the ones at the ends of closed interval. This generalization, as a matter of fact, is the basis of “extreme value theorem”.

Extreme value theorem

The extreme value theorem of continuous function guarantees existence of minimum and maximum values in a closed interval. Mathematically, if f(x) is a continuous function in the closed interval [a,b], then there exists f(l) ≤ f(x) and f(g) ≥ f(x) such that f(l) is global minimum and f(g) is global maximum of function.

As discussed earlier, there at least exists a pair of minimum and maximum at the end points. There may be more extremums depending on the nature of graph in the interval.

Range of function

If a function is continuous, then least i.e. global minimum, “A” and greatest i.e. global maximum, “B”, in the domain of function correspond to the end values specifying the range of function. The range of the function is :

$$[A, B]$$

If function is not continuous or if function can not assume certain values, then we need to suitably analyze function and modify the range given above. We shall discuss application of the concept of least and greatest values to determine range of function in a separate module.

Determining minimum and maximum values

There are three cases for determining minimum and maximum values. However, we should clearly underline that these methods give us relative minimum and relative maximum values – which may or may not be the greatest (global) or least (global) values. We need to interpret minimum and maximum in the context of specified domain to ascertain whether minimum and maximum are least and greatest respectively or not?

(i) function is differentiable in the domain of function.

(ii) function is continuous in the domain of function

(iii) function is discontinuous at certain points in the domain of function.

Function is differentiable

The derivative of function exists for all values of x in the domain. In this case, we follow the algorithm given here (without proof- its proof is based on Taylor’s expansion) :

1: Determine first derivative.

2: Equate derivative to zero.

3: Solve equation obtained in the step 2 for x.

4: If there is no real solution of equation, then function has no minimum or maximum.

5: If there is real solution of equation, then determine second derivative. Put root values in the expression of second derivative one after another and see whether second derivative is non-zero. If second order derivative is positive non-zero, then function is minimum at that root value. On the other hand, if second order derivative is negative non-zero, then function is maximum at that root value. We should note that these conclusions are valid for all higher even derivatives, which we might need to evaluate.

6: If second derivative is zero for any root value, then proceed to determine third derivative. If at any root value (which has not been decided in earlier step) third order derivative is non-zero, then function has no minimum or maximum at that root value. We should note that this conclusion is valid for all higher odd derivatives, which we might need to evaluate.

7: Continue with higher order even and odd derivatives till all root values are evaluated for minimum and maximum.

Example:
Problem : Determine minimum and maximum values of function :

$$f(x) = \frac{x^3}{3} - x$$

Solution : Differentiating with respect x, we have :

$$\Rightarrow f'(x) = \frac{1}{3} X 3x^2 - 1 = x^2 - 1 = (x-1)(x+1)$$

The roots of the corresponding equation are -1 and 1. Now, differentiating with respect to x again,

$$\Rightarrow f''(x) = 2x$$

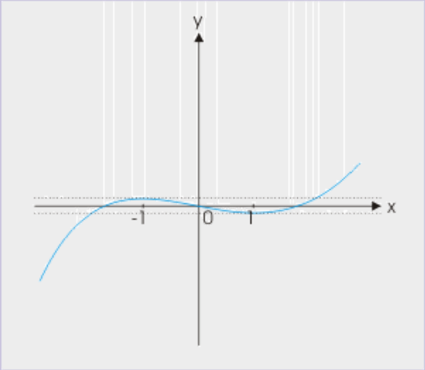
Putting, $x = -1, \quad f''(x) = -2 \quad < 0$. Hence, function has maximum value at x=-1.

$$\text{Maximum value} = \frac{(-1)^3}{3} - (-1) = -\frac{1}{3} + 1 = \frac{2}{3}$$

Putting, $x = 1, \quad f''(x) = 2 > 0$. Hence, function has minimum value at x=1.

$$\text{Minimum value} = \frac{(1)^3}{3} - 1 = \frac{1}{3} - 1 = -\frac{2}{3}$$

Graph of function



Minimum and maximum values of function.

Example:
Problem : Determine minimum and maximum values of function :

$$2x^3 - 9x^2 + 12x - 11$$

Solution : Differentiating with respect x, we have :

$$\Rightarrow f\prime(x) = 6x^2 - 18x + 12 = 6\left(x^2 - 3x + 2\right) = 6(x-1)(x-2)$$

The roots are 1 and 2. Now, differentiating with respect to x again,

$$\Rightarrow f\prime\prime(x) = 12x - 18$$

Putting, $x = 1, \quad f\prime\prime(x) = -6 \quad < 0$. Hence, function has maximum value at x=1.

$$\text{Maximum value} = 2x^3 - 9x^2 + 12x - 10 = 2X1^3 - 9X1^2 + 12X1 - 11 = -6$$

Putting, $x = 2, \quad f\prime\prime(x) = 6 \quad > 0$. Hence, function has minimum value at x=2.

$\text{Minimum value} = 2x^3 - 9x^2 + 12x - 11 = 2X2^3 - 9X2^2 + 12X2 - 11 = 16 - 36 + 24 - 11 = -7$

Example:
Problem : Determine minimum and maximum values of function :

$$f(x) = x + \frac{1}{x}$$

Solution :

$$\Rightarrow f\prime(x) = 1 - \frac{1}{x^2}$$

$$\Rightarrow f\prime(x) = \frac{x^2 - 1}{x^2}$$

The roots are -1 and 1. Now, differentiating with respect to x again,

$$\Rightarrow f\prime\prime(x) = \frac{2}{x^3}$$

At $x = 1, \quad f\prime\prime(x) = 2 \quad > 0$. Function has minimum value of 2 at x=1. At $x = -1, \quad f\prime\prime(x) = -2 \quad < 0$. Function has maximum value of -2 at x=-1. Note that minimum value is greater than maximum value. It is possible as function is not defined at x=0. A maximum of smaller value exist left to it and a minimum of higher value exist to the right of it.

Example:

Problem : Determine minimum and maximum values of function :

$$f(x) = x^5 - 5x^4 + 5x^3 - 5$$

Solution : Differentiating with respect x, we have :

$$\Rightarrow f\prime(x) = 5x^4 - 20x^3 + 15x^2$$

Equating to zero, we have :

$$5x^4 - 20x^3 + 15x^2 = 0$$

$$\Rightarrow x^4 - 4x^3 + 3x^2 = 0$$

$$\Rightarrow x^2\left(x^2 - 4x + 3\right) = 0$$

$$\Rightarrow x^2(x - 1)(x - 3) = 0$$

The roots are 0, 1 and 3. Now, differentiating with respect to x again,

$$\Rightarrow f\prime\prime(x) = 20x^3 - 60x^2 + 30x$$

Putting, $x = 0, \quad f\prime\prime(x) = 0$. We need to differentiate again to evaluate this point. Putting, $x = 1, \quad f\prime\prime(x) = -10 \quad < 0$. Hence, function has maximum value at x=1,

$$\text{Maximum value} = x^5 - 5x^4 + 5x^3 - 1 = 15 - 5X14 + 5X13 - 5 = 1 - 5 + 5 - 5 = -4$$

Putting, $x = 3, \quad f\prime\prime(x) = 90 \quad > 0$. Hence, function has minimum value at x=3.

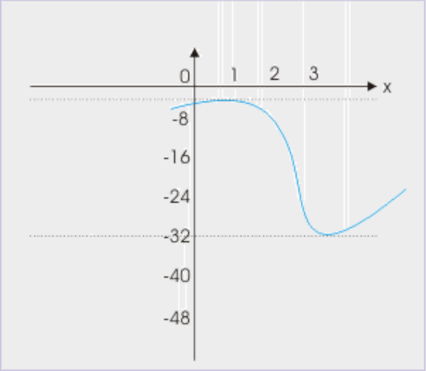
$$\text{Minimum value} = x^5 - 5x^4 + 5x^3 - 1 = 3^5 - 5X3^4 + 5X3^3 - 5 = 243 - 5X81 + 5X27 - 5 = -32$$

In order to determine nature at point x=0, we differentiate again,

$$\Rightarrow f\prime\prime\prime(x) = 60x^2 - 120x + 30$$

Putting, $x = 0, \quad f\prime\prime\prime(x) = 30 \quad > 0$. Hence, function has neither minimum nor maximum value at x=0.

Graph of function



Minimum and maximum values of function.

Function is continuous

We know that if function is continuous in an interval, then pair of minimum and maximum occur alternatively in any order. We shall use this fact to determine minimum and maximum. This technique being applicable to continuous function allows us to analyze even piece-wise defined functions. A continuous function may or may not be differentiable. For example, we can theoretically draw a modulus function without lifting pen. As such, it is a continuous function. However, it is not differentiable at x=0 where we can not draw a tangent.

In this case, we follow the algorithm given here:

1: Determine first derivative.

2: Draw sign diagram of first derivative. Note that it is slightly a different step than the equivalent step given for earlier case. Here, we are required to draw sign diagram **-** not the roots of first derivative equation.

3: If function is decreasing to the left and increasing to the right of a critical point of sign diagram, then function value at that point is minimum.

4: If function is increasing to the left and decreasing to the right of a critical point of sign diagram, then function value at that point is maximum.

Note : We can use this technique to determine minimum and maximum for function with undefined points as well. We shall illustrate this for a case of rational function in the examples given here.

Example:
Problem : Find minimum value of modulus function
Solution : Modulus function is defined as :
`| x ; x≥0 f(x) = | | -x ; x<0`
For x>0,

$$f(x) = x$$

$$\Rightarrow f\prime(x) = 1 \quad > 0$$

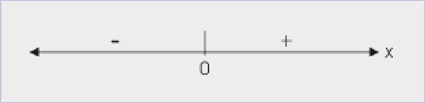
Since first derivative is positive, given function is increasing function for x>0.
For x<0,

$$f(x) = -x$$

$$\Rightarrow f\prime(x) = -1 \quad < 0$$

Since first derivative is negative, given function is decreasing function for x<0. Overall sign diagram of modulus function is as shown here :

Sign scheme



Sign scheme

At x=0, the function is decreasing to its left and increasing to its right. It means function has minimum at x=0.

$$\text{Minimum value} = 0$$

Note that minimum value in this case is also least value as there is only one minimum in the entire domain. Hence, minimum at x=0 is global minimum.

Example:

Problem : The function $y = a\log_e x + bx^2 + x$ has exteme values at x=-1 and x=2. Find values of “a” and “b”.

Solution : Here extreme values (maximum or minimum) are given. We know that first derivative is zero at extreme values. Now,

$$\Rightarrow y\prime = aX\frac{1}{x} + 2bx + 1$$

At x =-1,

$$\Rightarrow y\prime = aX\frac{1}{-1} + 2bX - 1 + 1 = 0$$

$$\Rightarrow -a - 2b + 1 = 0$$

At x = 2,

$$\Rightarrow y\prime = aX\frac{1}{2} + 2bX2 + 1 = 0$$

$$\Rightarrow \frac{a}{2} + 4b + 1 = 0$$

$$\Rightarrow a + 8b + 2 = 0$$

Solving two simultaneous equations,

$$\Rightarrow a = 2$$

$$\Rightarrow b = -\frac{1}{2}$$

Example:

Problem : Find maximum and minimum values of function :

$$y = \frac{x^2 - 7x + 6}{x - 10}$$

Solution : This function is not defined for x=10. The function is continuous except at this point. Thus, minimum and maximum obtained do not belong to a continuous domain.

$$\Rightarrow y\prime = \frac{(x-10)(2x-7) - \left(x^2 - 7x + 6\right)}{(x-10)^2} = \frac{2x^2 - 27x + 70 - x^2 + 7x - 6}{(x-10)^2}$$

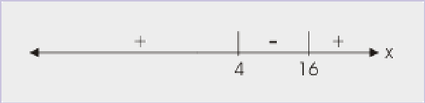
$$\Rightarrow y\prime = \frac{x^2 - 20x + 64}{(x-10)^2}$$

Now denominator is a positive number for all x. Thus, sign diagram of first derivative is same as that of numerator. In order to draw sign diagram, we need to factorize numerator.

$$\Rightarrow x^2 - 20x + 64 = (x-4)(x-16)$$

Hence, critical points are 4 and 16. The sign diagram is as shown in the figure. At x=4, the function is increasing to its left and decreasing to its right. It means function has maximum at x=4.

Sign scheme



Sign scheme

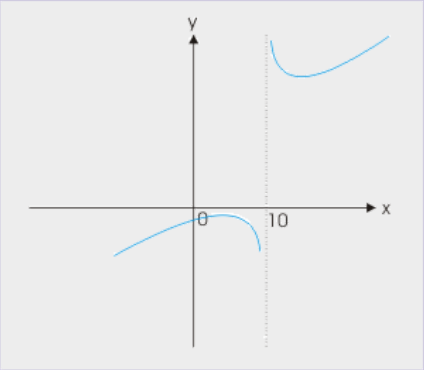
$$\text{Maximum value} = \frac{x^2 - 7x + 6}{x - 10} = \frac{4^2 - 7X4 + 6}{4 - 10} = 1$$

At x=16, the function is decreasing to its left and increasing to its right. It means function has minimum at x=16.

$$\text{Minimum value} = \frac{x^2 - 7x + 6}{x - 10} = \frac{16^2 - 7X16 + 6}{16 - 10} = 25$$

Note that minimum value is greater than maximum value.

Graph of function



The minimum and maximum values of a function.

Exercise

Exercise:

Problem: Determine maximum value of function :

$$f(x) = \left(\frac{1}{x}\right)^x$$

Solution:

The function is not defined for x=0. The function of the form x^y is defined for x>0. Comparing,

$$\Rightarrow \frac{1}{x} > 0$$

The critical point of this rational inequality is zero. The rational function 1/x is positive for x>0. Thus, domain of given function is x>0. In order to differentiate this function, we need to take logarithm. Let,

$$y = \left(\frac{1}{x}\right)^x$$

$$\Rightarrow \log_e y = x \log_e \left(\frac{1}{x}\right) = x(\log{}_e 1 - \log_e x) = -x \log_e x$$

Differentiating with respect to x, we have :

$$\Rightarrow \frac{1}{y} X \frac{dy}{dx} = -\log{}_e(x) - xX\frac{1}{x} = -(1 + \log{}_e x)$$

$$\Rightarrow \frac{dy}{dx} = -y(1 + \log_e x) = -\left(\frac{1}{x}\right)^x (1 + \log_e x)$$

In order to determine sign diagram of first derivative, we equate it to zero.

$$-\left(\frac{1}{x}\right)^{x}(1+\log_{e}x)=0$$

Now, $(1/x)^x > 0 \quad \text{as} \quad x > 0$. Hence sign diagram of first derivative is same as that of $-(1+\log_e x)$:

$$\Rightarrow -(1+\log_{e}x)=0$$

$$\Rightarrow \log_{e}x=-1$$

$$\Rightarrow x=e^{-1}=\frac{1}{e}$$

The expression 1/e is less than 1. We put x=1 to test the sign of right side. At x=1,

$$\Rightarrow -(1+\log_{e}x)=-(1+\log_e 1)=-(1+0)=-1 \quad < 0$$

This means function is increasing in interval (0,1/e] and decreasing in [1/e, ∞). Thus, function has maximum at x=1/e.

$$\text{Maximum value}=(1/x)^x=\left(\frac{1}{\frac{1}{e}}\right)^{\frac{1}{e}}=(e)^{1/e}$$

Note that this maximum value is greatest value as there is only one maximum in the domain of function.

Least and greatest function values

Least and greatest function values are characterizing aspects of a function. In particular, they allow us to determine range of function if function is continuous. Determination of these values, however, is not straight forward as there may be large numbers of minimum and maximum through out the domain of the function. It is difficult to say which of these are least or greatest of all. Two things simplify our analysis : (i) domain of investigation is finite and (ii) function is monotonic in sub-intervals within domain.

The study of least and greatest function values in this module is targeted to determine range of function. If “A” and “B” be the least and greatest values of a continuous function in a finite interval, then range of the function is given by :

$$\text{Range} = [\text{least, greatest}] = [\text{a,b}]$$

We should note that determining range is a comparatively more difficult proposition than determining domain. Recall that we need to solve given function for x to determine range. This solution, however, is not always explicit. As such, we may be stuck with problem of finding range of more complex functions – particularly those, which involves transcendental functions.

Further, we need to underline one important aspect, while evaluating range of a composite function. Range of a composite function is evaluated from inside to outside. This means that we need to evaluate innermost function and then the one outside it. This is an opposite order of evaluation with respect to domain which is evaluated from outside to inside. We shall highlight these aspects while working with examples.

In the following sections, we discuss various context of least and greatest values.

Standard functions

We are familiar with the least value, greatest value and range of the most standard functions of all origin. Consider constant, identity, reciprocal, modulus, greatest integer, least integer, fraction part, trigonometric, inverse trigonometric, exponential and logarithmic functions. All these functions have been described in detail and we know their properties with respect to least and greatest values and also the range. Greatest value of sine function, for example, is 1. On the other hand, exponential and logarithmic functions etc. neither have minimum (therefore least value) nor maximum (therefore greatest value). However, these functions have least and greatest in finite interval in accordance with mean value theorem.

In case, the function can be reduced to the standard forms having least and greatest values, then it is possible to know its range. In the example, we consider one such trigonometric function.

Example:
Problem : Find the range of the continuous function given by :

$$f(x) = a\cos x + b\sin x$$

where “a” and “b” are constants.
Solution : Here, given function is addition of two trigonometric functions. As we know least and greatest values of sine and cosine functions, we shall attempt to reduce given function in terms of either sine or cosine function (note that the algorithm for reducing addition of sine and cosine functions as presented here is a standard algorithm. We should also note that this algorithm, as a matter of fact, is used in analyzing superposition principle of waves) :

$$\text{Let} \quad a = r\cos\alpha \quad \text{and} \quad b = r\sin\alpha$$

where,

$$r = \sqrt{(a^2 + b^2)}$$

Substituting in the given function, we have :

$$f(x) = r\cos\alpha\cos x + r\sin\alpha\sin x = r\cos(x-\alpha) = \sqrt{(a^2+b^2)}\cos(x-\alpha)$$

We know that minimum and maximum values of cosine function are "-1" and "1" respectively. Hence,

$$f(x)_{\text{min}} = -\sqrt{(a^2+b^2)}$$

$$f(x)_{\text{max}} = \sqrt{(a^2+b^2)}$$

Therefore, range of the given function is :

$$\text{Range} = \left[-\sqrt{(a^2+b^2)}, \sqrt{(a^2+b^2)}\right]$$

Quadratic functions

The graph of quadratic function is known. If coefficient of second power term in the quadratic expression is positive, then function has no maximum and hence no greatest value, whereas its least value is given by “–D/4a”, where D is discriminant of corresponding quadratic equation and “a” is the coefficient of second power term. For x becoming large numbers, function value becomes very large positive value. Therefore, the range of function is given as :

$$\text{Range} = \left[-\frac{D}{4\text{a}}, \infty\right)$$

On the other hand, if coefficient of highest power term in the quadratic expression is negative, then greatest value of function is given by “–D/4a”, whereas there is no minimum and hence no least value. For x becoming large numbers, function value becomes very large negative value. Therefore, the range of function is given as :

$$\text{Range} = \left(-\infty, -\frac{D}{4\text{a}}\right]$$

We have already discussed these aspects in the module earlier and as such will not elaborate here again.

Monotonic function of same nature

Some functions have same monotonic nature through out domain. Linear polynomial, for example, is either strictly increasing or decreasing. There is no minimum or maximum value. Recall that minimum and maximum values occur between a pair of increasing and decreasing function values. Similar is the case with exponential and logarithmic functions. The ranges of exponential and logarithmic functions are open intervals. Since these functions are strictly increasing or decreasing though out their domain intervals, there is no minimum or maximum. As such, there is no least or greatest values of function.

However, if we investigate these functions in a finite interval, then function values at the boundary of closed interval are the least and greatest values in that interval. Further, functions, which are not singularly monotonic, can also be singly monotonic in a suitably selected interval. For example, sine function is a strictly increasing function in the interval [-π/2, π/2]. Thus, there are two possibilities :

1: Function is monotonic in one type without minimum or maximum function values. In any finite closed interval, however, the function has least and greatest values, which correspond to function values at interval ends.

2: A function is not monotonic of one type, but is rendered monotonic in suitably chosen finite closed interval. The function has least and greatest values, which correspond to function values at interval ends.

We shall use these properties to determine range of a function. In order to understand application of these concepts, we need to read each step of the examples as given here carefully :

Example:
Problem : Let $A = \left[\frac{\pi}{6}, \frac{\pi}{3}\right]$. If a continuous function is defined as :

$$f(x) = \cos x - x(1 + x)$$

Find range of the function.
Solution : “A” is an interval. We are required to find range of function. In order to find range, however, we would need to determine least and greatest function values. Towards this, we need to know the nature of function in the interval. For this, we find derivative of function and determine its sign. This will enable us to know whether the function is increasing or decreasing. Now,

$$f\prime(x) = -\sin x - 1 - 2x$$

The terms “sinx” and “2x” are positive for values of “x” in the interval specified by “A”. Therefore, we can conclude that the derivative is negative. It means that the function is a decreasing function in the entire interval. The end values of the function correspond to least and greatest values of the function and as such represent the range of the function. Now,

$$\Rightarrow f(x)_{\text{max}} = f(\pi/6) = \cos\frac{\pi}{6} - \frac{\pi}{6} - \left(\frac{\pi}{6}\right)^2 = \frac{\sqrt{3}}{2} - \frac{\pi}{6} - \frac{\pi^2}{36}$$

$$\Rightarrow f(x)_{\text{min}} = f(\pi/3) = \cos\frac{\pi}{3} - \frac{\pi}{3} - \left(\frac{\pi}{3}\right)^2 = \frac{1}{2} - \frac{\pi}{3} - \frac{\pi^2}{9}$$

Hence,

$$\Rightarrow \text{Range} = f(A) = \left[\frac{1}{2} - \frac{\pi}{3} - \frac{\pi^2}{9}, \frac{\sqrt{3}}{2} - \frac{\pi}{6} - \frac{\pi^2}{36}\right]$$

Example:

Problem 5: A function is defined as :

$$f(x) = \frac{e^x}{1 + [x]}$$

If domain of the function is [0,∞), find the range of the function.

Solution :

Statement of the problem : The function contains greatest integer function in its denominator, whereas numerator of functon is an exponential function. Greatest integer function, GIF, is defined in discontinuous intervals of 1. It returns integral value equal to the starting value of discontinuous interval. The function, however, is continuous in sub-intervals of 1 i.e. in a finite interval. It is clear that function values and function nature will change in accordance with the values of GIF in sub-intervals of the domain. We need to determine nature of function in few of the initial sub-intervals and extrapolate the results to determine the range of function as required.

In order to determine nature of function, we write function and its derivative in the initial sub-intervals of the given domain. Note that domain of function starts with 0.

$$0 \le x < 1, \qquad f(x) = e^x, \qquad f'(x) = e^x$$

$$1 \le x < 2, \qquad f(x) = \frac{e^x}{2}, \qquad f'(x) = \frac{e^x}{2}$$

$$2 \le x < 3, \qquad f(x) = \frac{e^x}{3}, \qquad f'(x) = \frac{e^x}{3}$$

The derivatives in each interval are positive for values of “x” in the domain. It means that function is strictly increasing in each of the intervals. There is no minimum and maximum as monotonic nature of function does not change. The least and greatest values, therefore, correspond to end function values in each finite sub-intervals. Using these least and greatest values, we determine range of initial sub-intervals as given here :

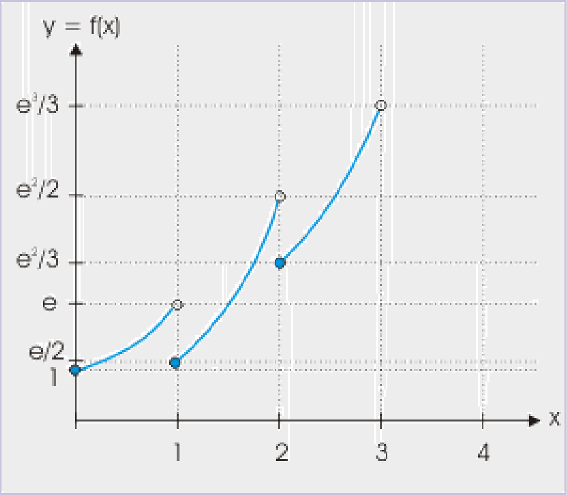
$$0 \le x < 1 \quad \Rightarrow e^0 \le y < e^1 \quad \Rightarrow 1 \le y < e$$

$$1 \le x < 2 \quad \Rightarrow \frac{e}{2} \le y < \frac{e^2}{2}$$

$$2 \le x < 3 \quad \Rightarrow \frac{e^2}{3} \le y < \frac{e^3}{3}$$

The plot of the function is drawn in the figure. Note that function values in adjacent intervals overlap each other, but exceeds the greatest value in the preceding intervals. Clearly, function value begins from “1” and continues towards infinity. Hence, range of the function is :

Range of the function



The range is divided in multiple overlapping intervals.

$$[1,\infty)$$

Function with finite intervals

A continuous function, in finite interval, has finite numbers of minimum and maximum values. Using this fact and the monotonic nature function in the finite interval, we can determine least and greatest values of function. If “A” and “B” be the least and greatest values then range of the function is given by :

$$\text{Range} = [A,B]$$

We need to compare minimum values, maximum values and end values to determine which are the least and greatest values. This aspect will be clear as we work out with the example here.

There are certain cases in which functions have exactly a pair of intervals characterized by opposite strict monotonic nature i.e. strictly increasing and strictly decreasing. In such cases, the point at which function changes its monotonic nature should be either minimum and maximum. This in turn means that function has exactly one minimum or maximum. Clearly, this minimum or maximum corresponds to least or greatest function value of the function in its domain.

Example:

Problem : Find range of continuous function :

$$f(x) = \sqrt{(x-1)} + \sqrt{(5-x)}$$

Solution : Given function is continuous. Also, it involves square root of polynomial. It means that function is defined in sub interval of real number set R. We, therefore, need to find the domain of the function first. Then we investigate the nature of the function in its domain and determine least (a) and greatest (b) function values.

For individual square roots to be real,

$$x - 1 \geq 0 \quad \Rightarrow x \geq= 1$$

$$5 - x \geq 0 \quad \Rightarrow x \leq 5$$

Hence, domain of the function is intersection of individual domains and is a finite interval [1,5]. Now, we differentiate the function with respect to independent variable to determine the nature of function :

$$\Rightarrow f\prime(x) = \frac{1}{2\sqrt{(x-1)}} - \frac{1}{2\sqrt{(5-x)}}$$

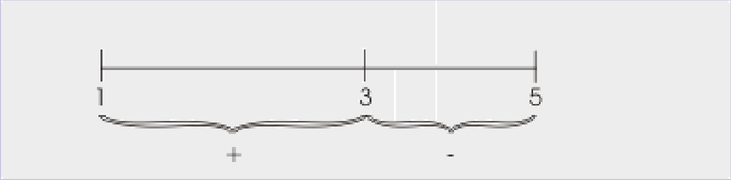
In order to draw sign scheme, we find the roots of the equation by putting f’(x) = 0 and solving for “x” as :

$$\Rightarrow 2\sqrt{(x-1)} = 2\sqrt{(5-x)}$$

$$\Rightarrow x - 1 = 5 - x \quad \Rightarrow 2x = 6 \quad \Rightarrow x = 3$$

For test value, x = 4, f’(x) is negative. Hence, corresponding sign scheme of f’(x) is as given here :

Sign scheme of derivative expression



The value at the point of change in the nature of function corresponds to greatest value of the function.

Since function is increasing to the left and decreasing to the right of root value, the function value at root value is maximum. Clearly, there is only one maximum. Thus, it is greatest value also. Further, either of two function values at end points is least value of the function in the domain . Now, function values at end points and root point are :

$$\Rightarrow f(1) = 0 + 2 = 2$$

$$\Rightarrow f(3) = \sqrt{2} + \sqrt{2} = 2\sqrt{2}$$

$$\Rightarrow f(5) = 2 + 0 = 2$$

Clearly, function at x = 3 is greatest value and either values at x = 1 and 5 is least. Hence, range of the function is :

$$\left[2,2\sqrt{2}\right]$$

Composite function

We shall use concept of least value, greatest value and monotonic nature to evaluate range of composite function. As pointed out in the beginning of the module, we shall proceed evaluation inside out i.e. proceeding from innermost to outermost.

Example:

Problem : A function is defined as :

$$f(x) = \left[\log_e\left\{\sin^{-1}\sqrt{(x^2+x+1)}\right\}\right]$$

The outermost square bracket, here, represents greatest integer function. Find range of the function.

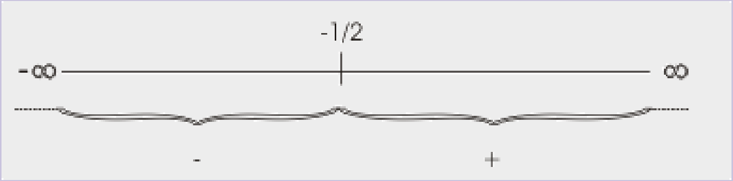
Solution : The argument of arcsine function is a square root. But we know that domain of arcsine function is [-1,1]. If we know the least or greatest value of quadratic expression within sqaure root, then we shall be able to futher narrow down the value of argument of arcsine function. The derivative of expression within square root is :

Now, the argument of arcsine function is a square root. The derivative of expression within square root is :

$$f\prime(x) = 2x + 1$$

The root of this expression when equated to zero is x = -1/2. The function value at test poin, x = 1 is 3, which is a positive number. Thus, corresponding sign scheme of the derivative expression is shown in the figure.

Sign scheme of derivative expression



The value at the point of change in the nature of function corresponds to the least value of the function.

The polynomial function is strictly decreasing to the left and strictly increasing to the right of root value. This means that function value at root point is minimum. Also, as there is only one minimum, this value is the least function value in the domain. Thus, least value of square root is :

$$\Rightarrow \sqrt{x^2+x+1} = \sqrt{\left\{\left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right) + 1\right\}} = \frac{\sqrt{3}}{2}$$

But, we know that the maximum value of the argument of arcsine function is “1”. Also, square root is a positive number. It means that :

$$\sqrt{x^2+x+1} \le 1$$

Thus, we conclude that the value of square root expression should lie in the interval given by :

$$\Rightarrow \frac{\sqrt{3}}{2} \le \sqrt{x^2+x+1} \le 1$$

We see that arcsine is an increasing function within the interval specified by domain of the function. Hence, we can take sine inverse of each term without any change in the direction of inequality :

$$\Rightarrow \sin^{-1}\frac{\sqrt{3}}{2} \le \sin^{-1}\sqrt{x^2+x+1} \le \sin^{-1}1$$

$$\Rightarrow \frac{\pi}{3} \le \sin^{-1}\sqrt{x^2+x+1} \le \frac{\pi}{2}$$

Logarithmic function is also an increasing function. Hence, we can take logarithm of each term without any change in the direction of inequality :

$$\log_{e}\frac{\pi}{3} \le \log_{e}\left(\sin^{-1}\sqrt{x^2+x+1}\right) \le \log_{e}\pi/2$$

Since $\log_e \pi/2 < 1$ and $\log_e \pi/3 > 0$, we can state the inequality as :

$$0 < \log_e\left(\sin^{-1}\sqrt{x^2+x+1}\right) < 1$$

We know that greatest integer function returns "0" in the interval [0,1). Hence,

$$f(x) = \left[\log_e\left(\sin^{-1}\sqrt{x_2+x+1}\right)\right] = 0$$

Clearly, function returns zero for all values of “x” in the domain of the function. Therefore, range of the function is :

$$\text{Range} = \{0\}$$

One-one and many-one functions (exercise)

A function is one – one function, if every pre-image(x) in domain is related to a distinct image (y) in the co-domain. Otherwise, function is many-one function.

Note : This module did not follow the treatment on the subject as we needed to know different types of real functions in the first place. Besides, there are additional methods to determine function types. In particular, the concept of monotonous functions (increasing or decreasing) can be used to determine whether a function is one-one or not.

Working rules

- Put $f(x_1) = f(x_2)$. Solve equation. If it yields $x_1 = x_2$, then function is one-one; otherwise not.
- Alternatively, draw plot of the given function. Draw a line parallel to x-axis such that it intersects as many points on the plot as possible. If it intersects the graph only at one point, then the function is one-one.
- Alternatively, put f(x) = 0. Solve f(x) = f(0) for "x" and see whether "x" is single valued for being one-one function.
- Alternatively, a function is an one-one function, if f(x) is a continuous function and is either increasing or decreasing function in the given domain.

Problem 1: A function $f : R \rightarrow R$ is given by :

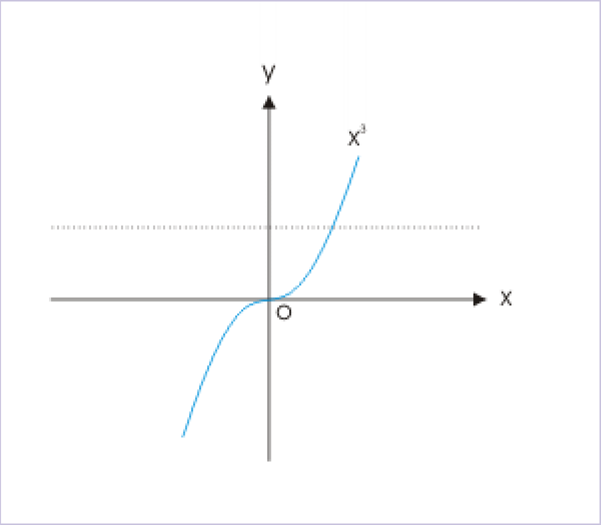
$$f(x) = x^3$$

Is the function one-one.

Solution :

Statement of the problem : Draw a line parallel to x-axis to intersect the plot of the function as many times as possible.

Function type



The line parallel to x-axis intersects function plot at one point.

We find that all lines drawn parallel to x-axis intersect the plot only once. Hence, the function is one-one.

Problem 2: A function $f : R \rightarrow R$ is given by :

$$f(x) = x^2$$

Is the function one-one.

Solution :

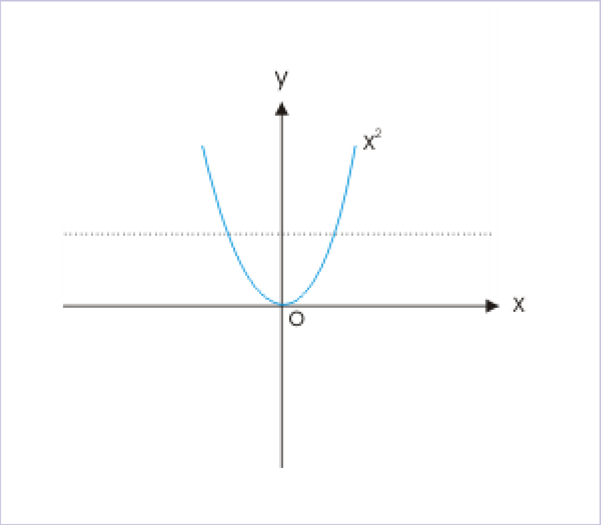
Statement of the problem : We can solve $f(x_1) = f(x_2)$ and see whether $x_1 = x_2$ to decide the function type.

$$\Rightarrow x_1^2 = x_2^2$$

$$x_1 = \pm x_2$$

We see that "x_1" is not exclusively equal to "x_2". Hence, given function is not one-one function, but many – one function. This conclusion is further emphasized by the intersection of a line parallel to x-axis, which intersects function plot at two points.

Function type



The line parallel to x-axis intersects function plot at two points.

Problem 3: A function $f : R \rightarrow R$ is given by :

$$f(x) = x|x|$$

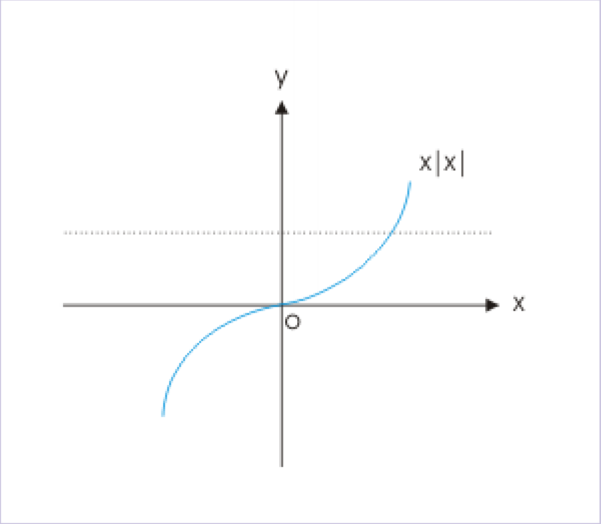
Is the function one-one.

Solution :

Statement of the problem : Draw the plot of the function and see intersection of a line parallel to x-axis.

We observe from its plot that there is no line parallel to x-axis, which intersects the functions more than once. Hence, function is one-one.

Function type

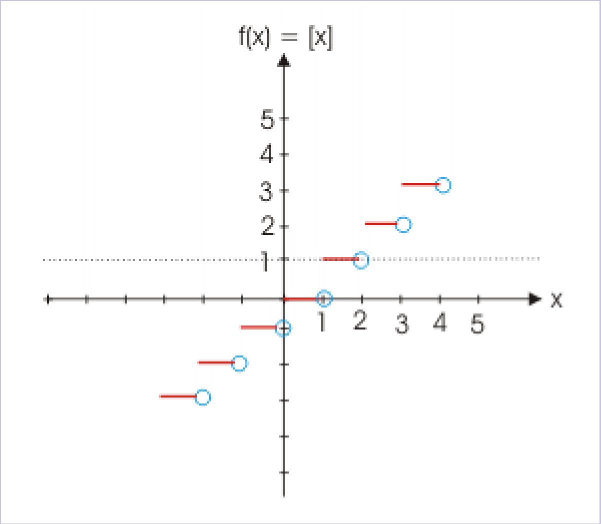


The line parallel to x-axis intersects function plot at one point.

Problem 4: Determine whether greatest integer function is one-one function.

Solution : Statement of the problem : Draw the plot of the function and see intersection of a line parallel to x-axis.

Function type



The line parallel to x-axis
intersects function plot at
infinite pointes.

We have drawn one such line at y = 1. We see that this function value is valid for an interval of “x” given by 1≤x<2. Hence, greatest integer function is not one-one, but many –one function.

Problem 5: A function $f : R \to R$ is given by :

$$f(x) = \frac{x^2 + 4x + 30}{x^2 - 8x + 18}$$

Is the function one-one.

Solution :

Statement of the problem : The given function is a rational function. We have to determine function type.

We evaluate function for x =0. If f(x)=f(0) equation yields multiple values of “x”, then function in not one-one. Here,

$$\Rightarrow f(0) = \frac{x^2 + 4x + 30}{x^2 - 8x + 18} = \frac{0^2 + 4X0 + 30}{0^2 - 8X0 + 18} = \frac{30}{18} = \frac{5}{3}$$

Now,

$$\Rightarrow f(x) = f(0)$$

$$\Rightarrow f(x) = \frac{x^2 + 4x + 30}{x^2 - 8x + 18} = \frac{5}{3}$$

$$\Rightarrow 3x^2 + 12x + 90 = 5x^2 - 40x + 90 \quad \Rightarrow 2x^2 - 52x = 0$$

$$\Rightarrow x = 0{,}26$$

We see that f(0) = f(26). It means pre-images are not related to distinct images. Thus, we conclude that function is not one-one, but many-one.

Problem 6 : A continuous function $f : R \to R$ is given by :

$$f(x) = \frac{x^2 + 4x + 30}{x^2 - 8x + 18}$$

Determine increasing or decreasing nature of the function and check whether function is an injection?

Solution :

Statement of the problem : The rational function is a continuous function. Hence, we can determine its increasing or decreasing nature in its domain by examining derivative of the function.

$$\Rightarrow f\prime(x) = \frac{\left(x^2 - 8x + 18\right)(2x + 4) - \left(x^2 + 4x + 30\right)(2x - 8)}{\left(x^2 - 8x + 18\right)^2}$$

$$\Rightarrow f\prime(x) = \frac{-12\left(x^2 + 2x - 26\right)}{\left(x^2 - 8x + 18\right)^2}$$

The denominator is a square of a quadratic expression, which evaluates to a positive number. On the other hand, the discreminant of the quadratic equation in the numerator is :

$$\Rightarrow D = (2)^2 - 4X1X(-26) = 4 + 104 = 108$$

It means that derivative has different signs in the domain interval. Therefore, the function is a combination of increasing and decreasing nature in different intervals composing domain. Thus, function is not monotonic in the domain interval. Hence, we conclude that function is not an injection.

Problem 7: A function

$$f : R \to R$$

is given by :

$$f(x) = \cos(3x + 2)$$

Is the function one-one.

Solution :

Statement of the problem : We solve $f(x_1) = f(x_2)$ and see whether $x_1 = x_2$ to decide the function type.

$$\cos(3x_1 + 2) = \cos(3x_2 + 2)$$

The basic solution is :

$$\Rightarrow 3x_1 + 2 = \pm(3x_2 + 2)$$

General solution is obtained by adding integral multiples of the period of function, which is “2π” for cosine function :

$$\Rightarrow 3x_1 + 2 = 2n\pi \pm (3x_2 + 2); \quad n \in Z$$

For n = 1,

$$\Rightarrow f(3x_1 + 2) = f\{2\pi \pm (3x_2 + 2)\}$$

Thus, multiple pre-images are related to same image. Hence, given function is not one-one.

We can easily interpret from the plots of different trigonometric functions that they are not one-one functions. However, they are one-one in the subset of their domain. Such is the case with other functions as well. They are generally not one-one, but may reduce to one-one in certain interval(s).

Problem 8 : A function $f : R \rightarrow R$ is given by :

$$f(x) = \frac{ax^2 + 6x - 8}{a + 6x - 8x^2}$$

Is the function one-one for a = 3?

Solution :

Statement of the problem : The given function is a rational function. Each of numerator and denominator functions is quadratic equation. In this case, solving f(a) = 0 for “x” reveals the nature of function for a =3.

$$\Rightarrow f(x) = \frac{3x^2 + 6x - 8}{3 + 6x - 8x^2} = 0$$

$$\Rightarrow 3x^2 + 6x - 8 = 0$$

$$\Rightarrow x = \frac{-6 \pm \sqrt{(36 - 4X3X - 8)}}{6} = -1 \pm \frac{\sqrt{33}}{3}$$

As “x” is not unique, the given function is not one-one function. We should emphasize here that solution of function when equated to zero is not a full proof method. In this particular case, it turns out that function value becomes zero for two values of “x”. In general, we should resort to techniques outlined in the beginning of the module to determine function type.

Inverse functions (exercise)

Working rules

A. Onto function

- A function is onto if every image in the co-domain has a pre-image in the domain set.
- Range = Co-domain

B. Bijection

- If a function is both one-one and onto, then the function is a bijection.

C. Inverse function

- Solve given function for “x”.
- Substitute “x” by inverse symbol “$f^{-1}(x)$” and “y” by “x” to find the rule of inverse function.
- Exchange the domain and range of the given function with that of inverse function.

Problem 1: A function $f : R \rightarrow R$ is given by :

$$f(x) = \sin(3x + 2)$$

Is the function surjective?

Solution :

Statement of the problem : Surjection is another name for onto function. Here, given function is a trigonometric function. We need to check whether its range is equal to co-domain.

According to question :

$$\text{Domain} = R$$

$$\text{Co-domain} = R$$

Now, let

$$y = f(x) = \sin(3x + 2)$$

Solving for "x", we have :

$$\Rightarrow 5x + 2 = \sin^{-1} y$$

$$\Rightarrow x = \frac{\sin^{-1} y - 2}{5}$$

The inverse trigonometric function is valid in its domain of [-1,1]. Hence, values of "y", for which "x" is real, is [-1,1]. Therefore,

$$\text{Range} = [-1, 1]$$

Clearly, range is not equal to co-domain. Thus, given function is not an onto function.

Problem 2: A function $f : R \rightarrow R$ is given by :

$$f(x) = \frac{ax^2 + 6a - 8}{a + 6x - 8x^2}$$

Find the interval of values of "a" for which function is onto.

Solution :

Statement of the problem : It is given that the rational function is valid for all values of "x" as its domain is "R". It is also given that co-domain of the function is "R". In order to find the interval of "a", we consider that the function is an onto function. This implies that range of the function is equal to co-domain of the function i.e. "R".

Let

$$y = \frac{ax^2 + 6a - 8}{a + 6x - 8x^2}$$

$$\Rightarrow ax^2 + 6x - 8 - ay - 6xy + 8x^2y = 0$$

Rearranging as quadratic equation in variable “x”, we have :

$$\Rightarrow (a + 8y)x^2 + 6(1 - y)x - (8 + ay) = 0$$

Since “x” is real, discriminant of the quadratic equation should be non-negative,

$$\Rightarrow 36(1 - y)^2 - 4X(a + 8y)X - (8 + ay) \geq 0$$

$$\Rightarrow 9(1 - y)^2 + (a + 8y)X(8 + ay) \geq 0$$

Expanding and rearranging as quadratic equation in variable “y”, we have :

$$\Rightarrow 9\left(1 + y^2 - 2y\right) + \left(8a + a^2y + 64y + 8ay^2\right) \geq 0$$

$$\Rightarrow (9 + 8a)y^2 + \left(a^2 + 46\right)y + (9 + 8a) \geq 0$$

But, we have assumed that function is an onto function. As such, range of the function is equal to co-domain i.e “R”. It means that above inequality holds for all real values of “y”. Now, remember (as explained for sign scheme of quadratic equation) that a quadratic expression evaluates to non-negative number if (i) coefficient of “y2” is a positive number and (ii) descreminant is a non-positive number. Hence,

$$\Rightarrow 9 + 8a > 0 \quad \Rightarrow a > -9/8$$

and

$$\Rightarrow \left(a^2 + 46\right)^2 - \left\{2(9 + 8a)\right\}^2 \leq 0$$

$$\Rightarrow \left(a^2 + 46 + 18 + 16a\right)\left(a^2 + 46 - 18 - 16a\right) \leq 0$$

$$\Rightarrow \left(a^2 + 16a + 64\right)\left(a^2 - 16a + 28\right) \leq 0$$

$$\Rightarrow (a + 8)^2(a - 2a - 14a + 28) \leq 0$$

$$(a+8)^2(a-2)(a-14) \le 0$$

$$\Rightarrow (a-2)(a-14) \le 0$$

$$\Rightarrow 2 \le a \le 14$$

The first condition earlier yielded as a > = 9/8. The required interval, therefore, is intersection of two intervals, which is [2,14].

Problem 3: Let “f” be a one-one function with domain [x,y,z] and range [1,2,3]. It is given that only one of three conditions given below are true and remaining two are false :

$$f(x) = 1; \quad f(y) \ne 1; \quad f(z) \ne 2$$

Determine $f^{-1}(1)$.

Solution :

Statement of the problem : With three conditions, there are three different possibilities for one being true and other two being false. We shall check each such possibility to see whether function is onto function. If the function is onto, then the function is a bijection as it is already given that function is an injection.

First combination :

$$\text{If} \quad f(x) = 1 \quad \text{is true, then} \quad f(x) = 1.$$

$$\text{If} \quad f(y) \ne 1 \quad \text{is false, then} \quad f(y) = 1.$$

$$\text{If} \quad f(z) \ne 2 \quad \text{is false, then} \quad f(z) = 2.$$

The function is clearly not one-one and hence not a bijection as f(x) = f(y).

Second combination :

$$\text{If} \quad f(x) = 1 \quad \text{is false, then} \quad f(x) = 2 \quad \text{or} \quad 3.$$

$$\text{If} \quad f(y) \ne 1 \quad \text{is true, then} \quad f(y) = 2 \quad \text{or} \quad 3.$$

$$\text{If} \quad f(z) \neq 2 \quad \text{is false, then} \quad f(z) = 2.$$

As f(z) = 2, f(x) = f(y) = 3. Again, the function is not one-one and hence not a bijection.

Third combination :

$$\text{If} \quad f(x) = 1 \quad \text{is false, then} \quad f(x) = 2 \quad \text{or} \quad 3.$$

$$\text{If} \quad f(y) \neq 1 \quad \text{is false, then} \quad f(y) = 1.$$

$$\text{If} \quad f(z) \neq 2 \quad \text{is true, then} \quad f(z) = 1 \quad \text{or} \quad 3.$$

As f(y) = 1, f(z) = 3 and f(x) = 2. In this case, we see that image and pre-image are related distinctly. The function is one-one and onto. Hence function is bijection for this combination. Also,

$$f^{-1}(1) = y$$

Problem 4: A function $f : R \rightarrow R$ is given by :

$$f(x) = x|x|$$

Find $f^{-1}(x)$.

Solution :

Statement of the problem : We first need to see that the given function is bijection. If the function is bijection, then we find the rule of corresponding inverse function following the algorithm given in the beginning of the module.

By definition of modulus function,

$$f(x) = xX - x = -x^2 \quad -\infty < x \leq 0$$

$$\Rightarrow f(x) = xXx = x^2 \quad 0 \leq x < \infty$$

We take the derivative of the function to check whether the function is an injection :

$$\Rightarrow f\prime(x) = -2x \quad -\infty < x \leq 0$$

$$\Rightarrow f\prime(x) = 2x \quad 0 \leq x < \infty$$

We see that derivative of function is non-negative number for all values of "x"

$$\Rightarrow f\prime(x) \geq 0$$

The equality holds only at a single point x = 0. Therefore, we conclude that function is an increasing function for all real values of "x". It means that given function is one-one function in the domain of "R". Now, we need to check whether function is onto function or not? For this, we find the range of the function. If range is "R", then range is equal to co-domain, which is also "R". For finding range of the function, we interpret the function, when "x" tends to become negative infinity or positive infinity (this is equivalent of taking limit).

$$\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} -x^2 = -\infty$$

and

$$\lim_{x \to \infty} f(x) = \lim_{x \to \infty} x^2 = \infty$$

The range of the function, therefore, is set of real numbers, "R". Thus, we conclude that given function is an onto function.

The given function is both one-one and onto function and hence it is a bijection.

In order to find inverse function, we first solve the equation for "x" as :

For the interval, $-\infty < x \leq 0$

$$\Rightarrow y = -x^2 \quad \Rightarrow x^2 = -y \quad \Rightarrow x = \pm\sqrt{(-y)}$$

But, we see that value of "x" is non-positive in the specified interval. Hence,

$$\Rightarrow x = -\sqrt{(-y)}$$

For the interval, $\Rightarrow 0 \le x < \infty$

$$\Rightarrow y = x^2 \quad \Rightarrow x^2 = y \quad \Rightarrow x = \pm\sqrt{(y)}$$

But, we see that value of "x" is non-negative in the specified interval. Hence,

$$\Rightarrow x = \sqrt{(y)}$$

Substituting "x" by "$f^{-1}(x)$" and "y" by "x", we have the rule of inverse function as :

$$f^{-1}(x) = -\sqrt{(-x)} \quad -\infty < x \le 0$$

and

$$f^{-1}(x) = \sqrt{(x)} \quad 0 \le x < \infty$$

Problem 5: A quadratic function is given as :

$$f(x) = x^2 + x + 1$$

Is the function invertible? If not, then find the intervals in which it is invertible. Also find corresponding inverse functions.

Solution :

Statement of the problem : The given function is a continuous function valid for all values of “x”. We need to analyze function to determine whether function is bijection.

Let "x_1" and "x_2" be two values. Then,

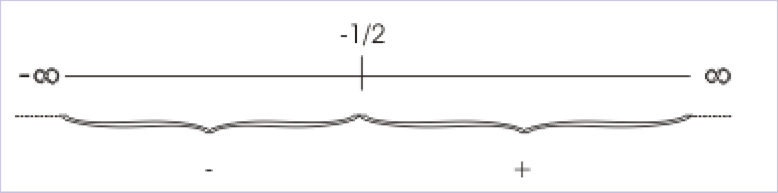
$$x_1^2 + x_1 + 1 = x_2^2 + x_2 + 1 \quad \Rightarrow x_1^2 = x_2^2 \quad \Rightarrow x_1 = \pm x_2$$

It means function is not one-one, but many one function. For determining whether function is onto function, we investigate the nature of its derivative,

$$f\prime(x) = 2x + 1$$

Its root, when equated to zero, is -1/2. Its sign scheme is shown in the figure. The function value at "x = -1/2" corresponds to least value of the function.

Sign scheme of derivative

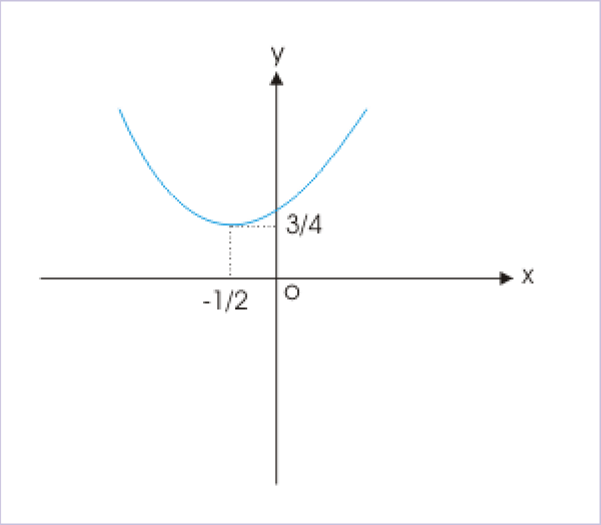


The function is strictly decreasing and increasing in two separate intervals.

$$\Rightarrow f\left(-\frac{1}{2}\right) = \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right) + 1 = \frac{1}{4} - \frac{1}{2} + 1 = \frac{3}{4}$$

From the figure it is clear that function is strictly decreasing in the interval (-∞,-1/2] and strictly increasing in the interval [-1/2, ∞). Further, we also observe from the graph as shown in the figure below that function is one-one in the individual interval.

Sign scheme of derivative



The function is strictly decreasing and increasing in two separate intervals.

Now, for determining inverse function, we solve the function for "x". Here :

$$y = f(x) = x^2 + x + 1$$

$$\Rightarrow x^2 + x + 1 - y = 0$$

Solving for “x”, we have :

$$\Rightarrow x = \frac{-1 \pm \sqrt{(4y-3)}}{2}$$

Substituting “x” by “$\Rightarrow f^{-1}(x)$” and “y” by “x”, we have the rule of inverse function as :

For interval $\Rightarrow \left(-\infty, -\frac{1}{2}\right]$

$$\Rightarrow f^{-1}(x) = \frac{-1 - \sqrt{(4x-3)}}{2}$$

For interval $\left[-\frac{1}{2},\infty\right)$

$$\Rightarrow f^{-1}(x)=\frac{-1+\sqrt{(4x-3)}}{2}$$

Problem 6: An onto function is given as :

$$f(x)=\log_a\left\{x+\sqrt{(x^2+1)}\right\},a>0,a\neq 1$$

Is the function invertible? If invertible, then find $f^{-1}(x)$.

Solution :

Statement of the problem : The given function is an onto function. This will be bijection i.e. invertible, if function is strictly monotonic (increasing or decreasing) so that function is one-one as well.

We see that $\sqrt{(x^2+1)}\geq 0$ for all values of “x”. Thus, argument of the logarithmic function is positive for all real values of “x”. It means that domain of the given function is “R”. In order to determine monotonic nature of the function, we differentiate given function in relevant intervals. Before differentiating, we need to convert the base from “a” to “e” as :

$$f(x)=\log e\left\{x+\sqrt{(x^2+1)}\right\}X\log_a e$$

Differentiating with respect to independent variable,

$$\Rightarrow f\prime(x)=\frac{\log_a e}{\left\{x+\sqrt{(x^2+1)}\right\}}X\quad 1+\frac{2x}{\left\{x+\sqrt{(x^2+1)}\right\}}$$

$$\Rightarrow f\prime(x)=\frac{\log_a e}{2\sqrt{(x^2+1)}}$$

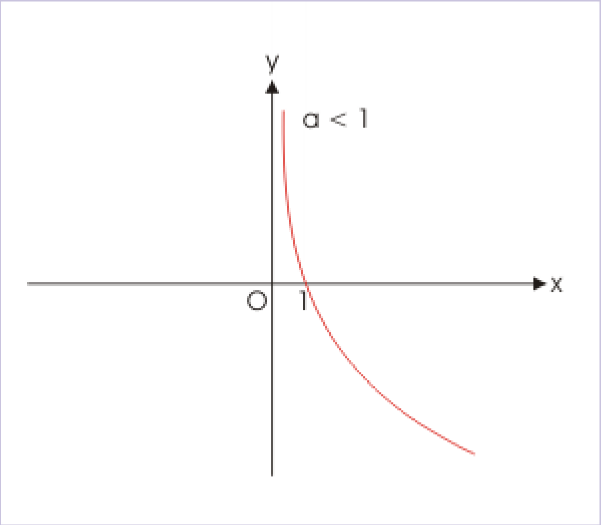
Again $\sqrt{(x^2+1)} \geq 1$ i.e. a positive number. Hence, sign of f'(x) is same as that of

$$\log_a e$$

.

For $0 < a < 1$

See the graph. Here x = e = 2.718281828.
Sign scheme of derivative



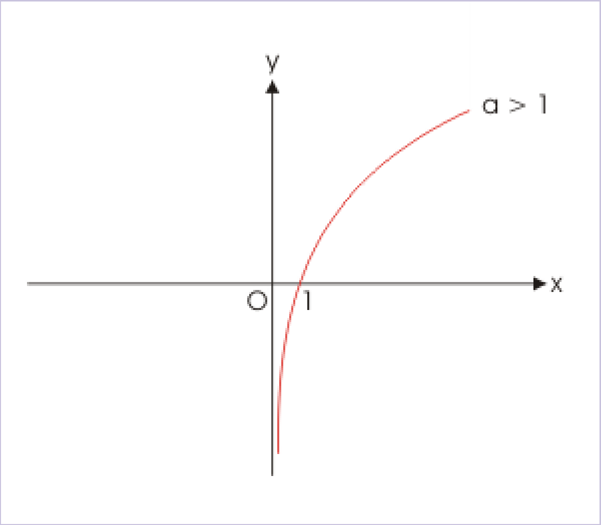
The function is strictly decreasing and increasing in two separate intervals.

$$\Rightarrow \log_a e < 0 \quad \Rightarrow f\prime(x) < 0$$

The function is a decreasing function in the domain i.e. "R"

For $a > 1$

See the graph. Here x = e = 2.718281828.
Plot of the function



The function is strictly decreasing and increasing in two separate intervals.

$$\Rightarrow \log_a e > 0 \quad \Rightarrow f\prime(x) > 0$$

The function is an increasing function in the domain i.e. “R”

We see that function is either increasing or decreasing for all real values of “x”. Thus, we conclude that function is strictly monotonic and function is on-one. This, in turn, means that function is bijection, which signifies that inverse of the function exists.

In order to find the inverse function, we write the logarithmic function in the equivalent exponential function as :

$$y = f(x) = \log_a\left\{x + \sqrt{(x^2+1)}\right\}$$

$$\Rightarrow a^y = x + \sqrt{(x^2+1)}$$

Also,

$$\Rightarrow a^{-y} = \frac{1}{a^y} = \frac{1}{x + \sqrt{(x^2+1)}}$$

$$\Rightarrow a^{-y} = \frac{\sqrt{(x^2+1)} - x}{\left\{x + \sqrt{(x^2+1)}\right\}\left\{\sqrt{(x^2+1)} - x\right\}}$$

$$\Rightarrow a^{-y} = \sqrt{(x^2+1)} - x$$

In order to solve for “x”, we subtract two equations :

$$\Rightarrow a^y - a^{-y} = x + \sqrt{(x^2+1)} - \sqrt{(x^2+1)} + x = 2x$$

$$\Rightarrow x = \frac{1}{2}\left(a^y - a^{-y}\right)$$

Substituting “x” by “$f^{-1}(x)$ ” and “y” by “x”, we have the rule of inverse function as :

$$\Rightarrow f^{-1}(x) = \frac{1}{2}\left(a^x - a^{-x}\right)$$

Problem 7: A function f : $[1,\infty) \rightarrow [1,\infty)$ is given by :

$$f(x) = 2^{x(x-1)}$$

Find $f^{-1}(x)$.

Solution :

Statement of the problem : In order to determine the nature of given function to be a bijection, we need to check whether the function is both one-one and onto?

To check for one-one function, we determine the derivative of the function as :

$$f(x) = 2^{x(x-1)}$$

Taking logarithm on both sides,

$$\log_e(f(x)) = x(x-1)\log_e 2$$

Taking derivative on either side of the equation, we have :

$$\frac{f\prime(x)}{f(x)} = (2x-1)\log_e 2 \quad \Rightarrow f\prime(x) = f(x)(2x-1)\log_e 2$$

Now, “x” lie in the interval [1,∞). Hence, f(x) is a positive number. Also, $\log_e 2$ is a positive number. Therefore,

$$f\prime(x) = \text{positive number}X(2x-1)$$

The root of the expression of f’(x), when equated to zero, is “0.5”. The derivative is positive in the interval between “1/2” and infinity. However, the domain of function begins at x = 1. We, therefore, conclude that the function is increasing in the interval given by [1,∞). The least value of the function is given as :

$$\Rightarrow f(1) = 2^{x(x-1)} = 2^0 = 1$$

We interpret the function, when “x” tends to become positive infinity (this is equivalent of taking limit).

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} 2^{x(x-1)} = 2^{\infty X(\infty-1)} = 2^{\infty X\infty} = 2^{\infty} = \infty$$

Thus, range of the given function is [1, ∞). Clearly,

Range = Co-domain = [1, ∞)

Therefore, function is onto and hence bijection. It means that function is invertible. Now, we solve the function for “x” for finding inverse function. Taking log on the base of “2” on either side of the equation,

$$\Rightarrow \log_2 y = \log_2 x(x-1) = x(x-1) = x^2 - x$$

$$\Rightarrow x^2 - x - \log{}_2 y = 0$$

$$\Rightarrow x = \frac{1 \pm \sqrt{(1 + 4X1X\log_2 y)}}{2} = \frac{1 \pm \sqrt{(1 + 4\log_2 y)}}{2}$$

But, we know that domain of the function is x >= 1. Also, $\sqrt{(1 + 4\log_2 y)} \geq 1$. It means that only positive sign in the expression is valid.

$$\Rightarrow x = \frac{1 + \sqrt{(1 + 4\log_2 y)}}{2}$$

Substituting “x” by “$f^{-1}(x)$” and “y” by “x”, we have the rule of inverse function as :

$$\Rightarrow f^{-1}(x) = \frac{1 + \sqrt{(1 + 4\log_2 x)}}{2}$$

Limits

Limit is a concept which aims to determine nature of function at a point. This point is infinitesimally close to a declared or test point say “a”. We investigate nature of function when independent variable approaches the test value “a” and is not at “a”. In the nutshell, we seek to estimate value of function at x=a from a point which is very close to it. Definitely, neither “x” reaches “a” nor f(x) reaches a particular value, say, L. Thus, important thing is to understand that limit denotes correspondence of independent and dependent variables very near but not at the point of estimation.

We should keep in mind while studying limit that it is an estimation based on the behavior of function at points very near to the test point. Limit answers the question : “what would be function value at the test point from its behavior at a point which is very close?”. In answering this question, limit considers the nature of function as described by function rule and by estimating value at test point from either direction. This estimate or projection may, however, fail to match actual function value at the test point, if there is a jump or sudden change in function value i.e. when function is discontinuous at the test point. It does not matter. An estimate (limit) remains or exists – if it can be estimated – irrespective of whether it matches function value or not and whether there is a function value at all at the test point or not?

Delta – epsilon definition

Idea here is to express nature of function near a point, however, close. We can do this by choosing two very small positive numbers delta (δ) and epsilon (∈). We say that limit of function f(x) is L at x = a, if “x” approaches very close to “a”, then f(x) approaches very close to L. This means simultaneous closeness :

$$L - \delta < f(x) < L + \delta \quad \text{for all x in} \quad a- \in < x < a+ \in$$

In modulus form :

$$|f(x) - L| < \delta \quad \text{for all x in} \quad |x - a| < \in$$

Limit of function is L, which may or may not be equal to value of function at x=a i.e. f(a). We shall discuss this aspect subsequently.

Notation

Limit of a function is denoted as :

$$\lim_{x \to a} f(x) = L$$

We should read this notation carefully. It is the “limit of function” which is "equal to" L - not the function value. As far as function is concerned it is approaching “L” and value of function, f(a), may or may not be equal to “L”.

Nature of function

Nature of function is not known by its value at a point. Rather, it is known by the value it is likely to have at a neighboring point. Here, we consider hypothetical set up in order to understand the concept. Our job is to find the approaching value which the function will have and which can be represented as “L”. This we do by determining function value a little to the right towards test point if we approach the test point from left. Similarly, we approach the test point from right by determining function value a little to the left towards test point. If we approach the same value from either side from a very close point, then we say that limit of function at test point is “L”.

Important to note here is that this approaching value of function, L, at the test point, x=a, may or may not be the function value f(a). We should understand that the mechanism of piece-wise function definition allows us to define any function value for any point in the domain of function. Further, if test point is a singularity of domain (a point where function is not defined), then there is no function value at the test point.

Left hand limit or left limit

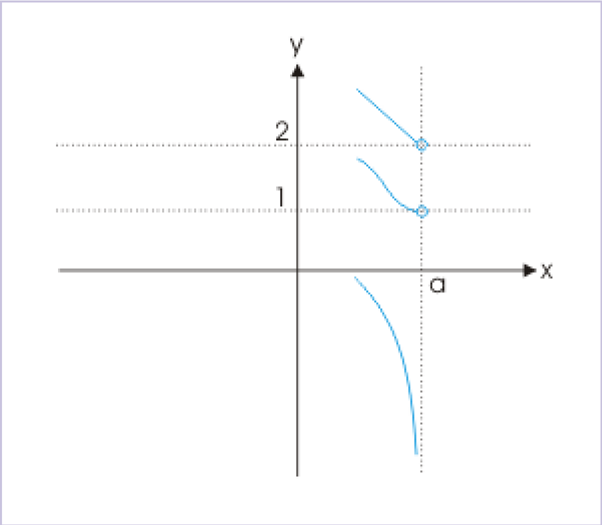
Left hand limit is an estimate of function value from a close point on the left of test point. It answer : what would be function value – not what is - at the test point as we approach to it from left? Symbolically, we represent this limit by putting a “minus” sign following test point “a” as “a-“.

$$\lim_{x \to a-} f(x) = L_l$$

In terms of delta – epsilon definition, we write :

$$L_l - \delta < f(x) < L_l + \delta \quad \text{for all x in} \quad a- \in < x < a$$

Left hand limit



Left hand limit

Graphically, we represent left limit by a curve which points towards limiting value from left terminating with an empty small circle at the test point. The empty circle denotes the limiting value. Since it is an estimate based on nature of graph – not actual function value, it is shown empty. In case, function value is equal to left limit, then circle is filled. If limit approaches infinity, then we show a graph with out terminating circle, approaching an asymptote towards either positive or negative infinity.

Right hand limit or right limit

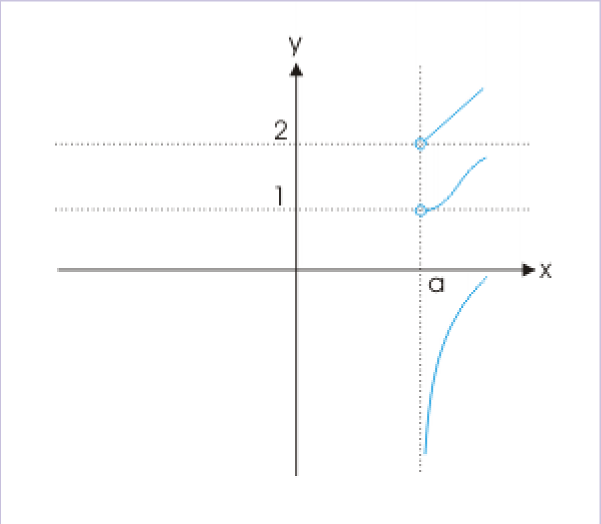
Right hand limit is an estimate of function value from a close point on right of test point. It asnwers : what would be function value – not what is - at the test point as we approach to it from right? Symbolically, we represent this limit by putting a “plus” sign following test point “a” as “a+“.

$$\lim_{x \to a+} f(x) = L_r$$

In terms of delta – epsilon definition, we write :

$$L_l - \delta < f(x) < L_l + \delta \quad \text{for all x in} \quad a < x < a + \in$$

Right hand limit



Right hand limit

Graphically, we represent right limit by a curve which points towards limiting value from right terminating with an empty small circle at the test point. If limit approaches infinity, then we show a graph with out terminating circle, approaching an asymptote towards either positive or negative infinity.

Limit at a point

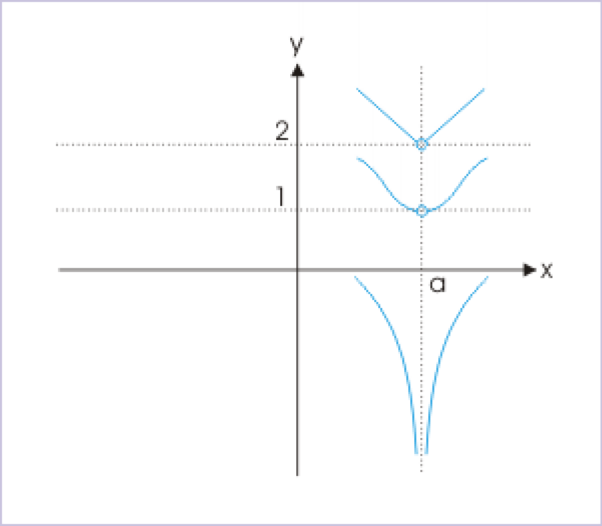
Limit is an estimate of function value from close points from either side of test point. If left and right limits approach same limiting value, then limit at the point exists and is equal to the common value. Clearly, if left and right limits are not equal, then we can not assign an unique value to the estimate. Clearly, limit of a function answers : what would be function value – not what is - at the test point as we approach to it from either direction? Symbolically, we represent this limit as :

$$\lim_{x \to a} f(x) = L_l = L_r = L$$

In terms of delta – epsilon definition, we write :

$$L - \delta < f(x) < L + \delta \quad \text{for all x in} \quad a - \in < x < a + \in$$

Limit at a point



Limit at a point

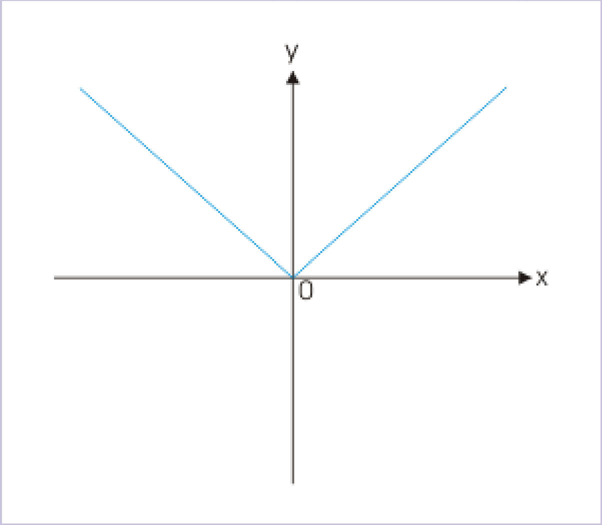
Graphically, we represent the limit by a pair of curves which point towards limiting value from left and right terminating with a common empty small circle at the test point. If limit approaches infinity, then we show a graph with out terminating circle, approaching an asymptote from either direction in the direction of either positive or negative infinity.

Limit and continuity

It has been emphasized that limit is an estimate of function value based on function rule at a point. This estimate is not function value. Function value is defined by the definition of function at that point. However, if function is continuous from the neighboring point to the test point, then limit should be equal to function value as well. Consider modulus function :

`| x ; x>0 f(x) = | 0 ; x=0 | -x ; x<0`

Modulus function



Modulus function

Clearly, at test point x=0,

$$L = f(0) = 0$$

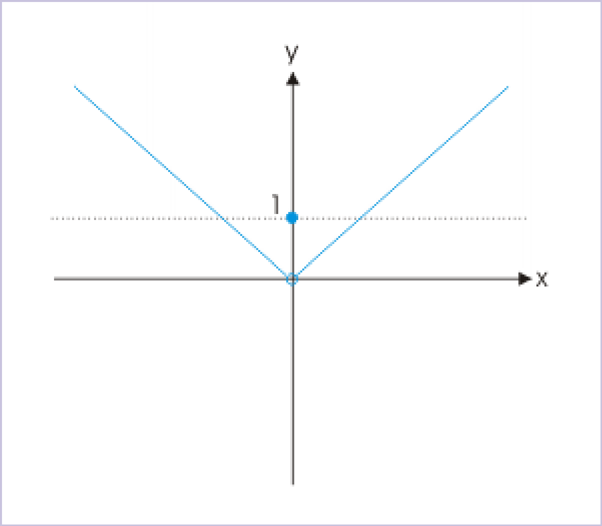
This is an important result which gives us a method to determine limit of a function. If function is continuous, simply put the test value into function definition. The value of function is limit of function at that point.

Limit and discontinuity

If function rule changes exactly at the test point, then limit of the function, L, and value of function, f(a), are not same. In order to clearly understand the implication of the statement about inequality of limit and function value, we consider a modification to the modulus function :

`| x; x>0 f(x) = | 1; x=0 | -x; x<0`

Modified modulus function



Modified modulus function

At test point x=0,

$$L = 0$$

$$f(0) = 1$$

$$L \neq f(0)$$

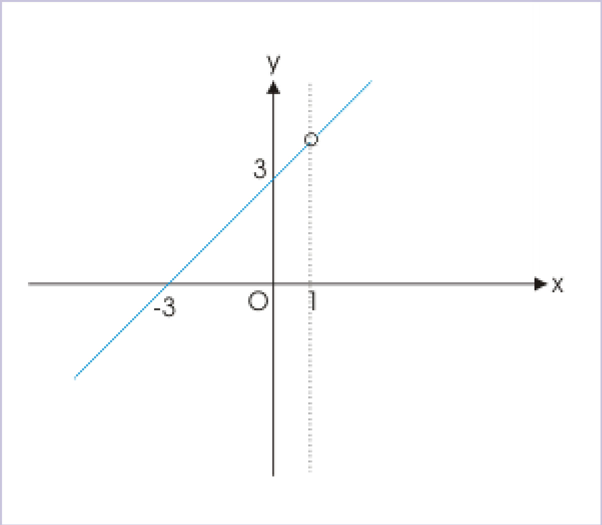
Therefore, limit of function exists at x=0 even though it is not equal to function value. This is an important result which gives us a method to determine limit of a piece-wise defined functions. We need to evaluate function from both left and right side. If limits are equal from both sides, then limit of function at test point is equal to either limit. However, if left and right limits are not equal then limit of function does not exist at the test point.

Limit and singularity

Singularity or exception point is a point where function is not defined. It is outside definition of function. However, function can point (or tend or approach) to a value at a point where it is not defined. Limit as we know estimate value from a close point where function exits and can project a value based on function definition at points very close to exception point. Consider limit of a rational function :

$$\lim_{x \to 1} \frac{(x-1)(x+3)}{(x-1)}$$

Rational function



Rational function

The singularity of function is obtained by setting denominator to zero. Thus, singularity exists at x=1. We want to know nature of function about this point. In other words, we want to know what would have been the value of function at this point had the function been defined there. For this, we need to evaluate left and right limit at this point. Graphically, there is a hole in the graph of the function. How can we estimate value of function at a point if it is not defined there ?

We keep linear factor in the denominator to know singularity. Extrapolating value at the singularity is a reverse process. We need to calculate function value in the neighborhood, where function is defined. For this, we require to remove linear factor from the denominator. Canceling out (x-1) from both numerator and denominator, we have :

$$\lim_{x\to 1}\frac{(x-1)(x+3)}{(x-1)} = \lim_{x\to 1}(x+3) = 4$$

Thus, function is not defined at x=1, but its limit at the point is 4. This means that nature of function in its immediate vicinity is such that the function should have attained a value of 4 had it been estimated on the basis of nature of function in the neighborhood.

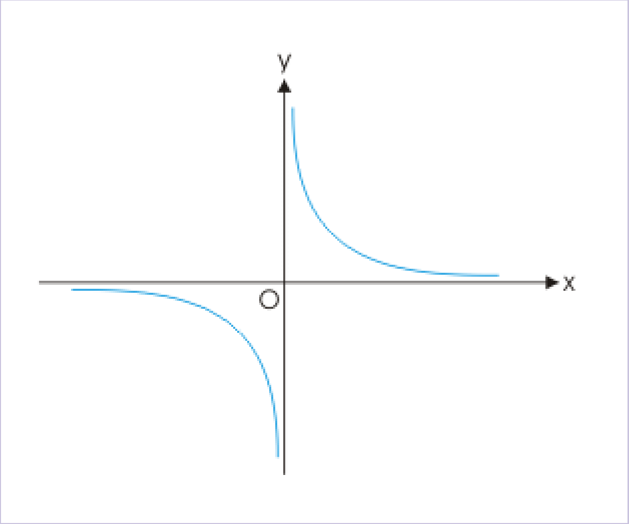
This is again an important result which gives us a method to determine limit of a function, when function is not defined at certain point or has other indeterminate or meaningless forms. We need to simplify function expression till we get a form which can be evaluated.

Let us now consider reciprocal function :

$$f(x) = \frac{1}{x}$$

Its singularity is obtained by setting denominator to zero. Thus, singularity exists x=0 for the reciprocal function. As such, domain of function is R-{0}. In order to know the function, we need to know nature of function in the vicinity of undefined point. We can do this by evaluating limit on either side of the singularity.

Reciprocal function



Reciprocal function

$$\lim_{x \to 0-} \frac{1}{x} = -\infty$$

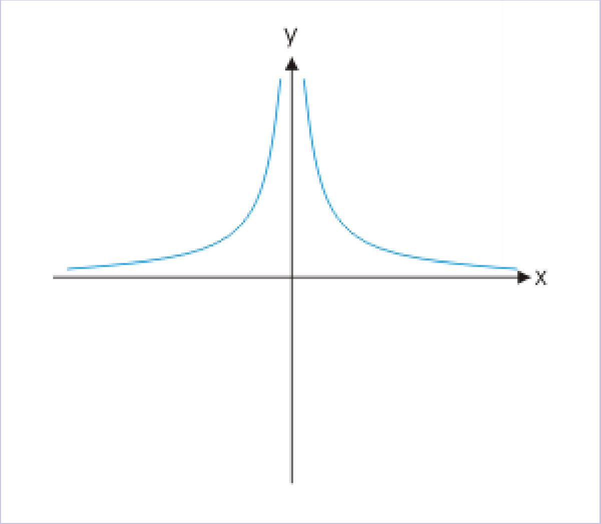
$$\lim_{x \to 0+} \frac{1}{x} = \infty$$

Important point to underscore here is that limiting values of x or f(x) as infinity is a valid estimates. To be more explicit, value of function can approach infinity as limiting value. In this case, left and right limits are not same. Therefore, limit of function does not exist at exception point x=0. In order to explore limit at exception point, we consider the case of modulus of reciprocal function. In this case also, function is not defined at x=0. But, for x<0; |x| = -x and for x>0; |x| = x. Hence, left and right limits are :

$$\lim_{x \to 0-} \left|\frac{1}{x}\right| = \lim_{x \to 0-} -\frac{1}{x} = \infty$$

$$\lim_{x \to 0+} \left|\frac{1}{x}\right| = \lim_{x \to 0+} \frac{1}{x} = \infty$$

Modulus of reciprocal function



Modulus of reciprocal function

In this case, left and right limits are equal. Therefore, limit of function exist at exception point x=0 and it is given as :

$$\lim_{x \to 0} \left|\frac{1}{x}\right| = \infty$$

Limits and infinity

We have noted that limit of function can be positive or negative infinity to reflect the estimate that function value is expected to be either very large positive or negative value. It happens when a finite value is divided by a value which is exceedingly small. If the divisor is a exceedingly small negative value, then function approaches negative infinity and if the divisor is a exceedingly small positive value, then function approaches positive infinity. Similar intuitive limiting values involving infinity are given here :

(1) Let “a” be a finite real number.

$$\lim_{x\to\infty} \frac{a}{x} = 0$$

$$\lim_{x\to-\infty} \frac{a}{x} = 0$$

(2) Let “a” be a finite real number.

$$\lim_{x\to0-} \frac{a}{x} = -\infty$$

$$\lim_{x\to0+} \frac{a}{x} = \infty$$

(3) Let “a” be a finite real number.

$$\lim_{x\to\infty} ax = \infty; \quad a > 0$$

$$= 0; \quad a = 0$$

$$= -\infty; \quad a < 0$$

(4) Let “a” be a finite real number.

$$\lim_{x\to\infty} a^x = \infty; \quad a > 1$$

$$= 1; \quad a > 0$$

$$= -\infty; \quad 0 < a < 1$$

Indeterminate limit forms

The indeterminate limit form is also called meaningless form. There are seven such forms in total. We, however, need to be careful in interpreting these forms. The interpretation is most important part of evaluation of limit. For example, if we say that 0/0 is indeterminate limit form, then we mean that both numerator and denominator of function approach zero, but none are equal to zero. In the example below, both numerator and denominator approach to zero as x approaches 2 :

$$\lim_{x\to 2}\frac{(x^2-4)}{(x-2)}$$

As x approaches 2, both numerator and denominator approaches to zero. Therefore, the function expression is an indeterminate form 0/0. However, following is not an indeterminate form :

$$\lim_{x\to 0}\frac{0}{x}=\infty$$

In the limit given above, the numerator is zero (not approaches to zero), whereas denominator approaches to zero. Thus, the rational form is determinate form and approaches infinity and the limit is also infinity.

In addition to 0/0 indeterminate form, there are other indeterminate forms which needs to be converted to determinate form so that limit can be evaluated. The seven indeterminate forms are :

$$0/0,\quad \infty/\infty,\quad 0.\infty,\quad \infty-\infty,\quad 0^{0,}\quad \infty^{0,}\quad 1^{\infty}$$

Again, we emphasize to distinguish interpretation with each of the indeterminate limit forms given above. In order to clarify the point further, let us consider two limits :

$$\lim_{x\to \pi/2}\sin x^{\tan x}$$

$$\lim_{x\to \pi/2}1^{\tan x}=1$$

In the first case, the base is approaching 1 and exponent is approaching infinity. Hence, it is indeterminate limit form. In the second case, base is 1 – not approaching to 1. Hence, it is not an indeterminate form and is evaluated to 1.

Evaluation of limit

There are three distinct regimes based on the discussion as above. We evaluate limit in accordance with following algorithm :

1: The function is not in indeterminate function form.

We simply plug in the value of test point into the function. The function value is equal to the limit of function.

2: The function is in indeterminate function form.

We transform the functions into determinate form. There are many techniques to transform an indeterminate form. Rationalization, simplification etc are important means to change forms. Expansion series of transcendental functions are also helpful. Besides, there are forms whose limits are known. We attempt to structure given expression in those standard forms and then find the limit. Finally, there are specific algorithms depending on nature of function, which help to remove indeterminate form and find limit. We shall study specific techniques in specific context.

3: The function is piecewise defined.

Using two approaches outlined above, we determine left and right limit and see whether they are equal or not? If equal, then limit is equal to either of left and right limits.

Note : We shall divide evaluation of limit in separate categories for different function types. The evaluations of limits shall be dealt separately in detail. Here, we work with few fundamental limits only.

Example

Example:
Problem : Plot the graph of function and determine limit when x->1.

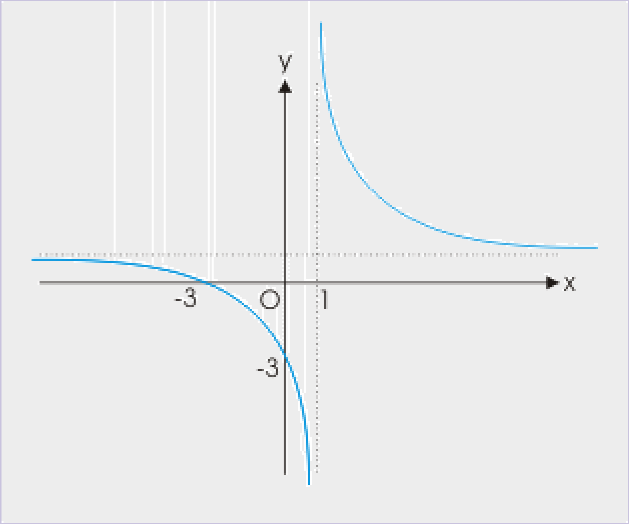
$$f(x) = \frac{x+3}{x-1}$$

Solution : Here, singularity exists at x=1. The function is not defined at this point. Rearranging, we have :

$$\frac{x+3}{x-1} = \frac{x-1+4}{x-1} = 1 + \frac{4}{x-1}$$

Here core graph is 4/x graph. We obtain graph of 4/(x-1) by shifting graph of 4/x towards right by 1 unit. In order to obtain the graph of given function, we shift the graph of 4/(x-1) up by 1 unit. We see that the function has one zero at x=-3. It has one asymptote at x=1. On the other hand, the y-intercept of function is -3.

Graph of rational function



Graph of rational function

Left hand limit is :

$$\Rightarrow \lim_{x \to 1-} 1 + \frac{4}{x-1} = -\infty$$

Right hand limit is :

$$\Rightarrow \lim_{x \to 1+} 1 + \frac{4}{x-1} = \infty$$

Since left and right hand limits are not equal, the limit of function does not exist as x approaches to 1.

Exercises

Exercise:

Problem: Determine limit :

$$\Rightarrow \lim_{n \to \infty} \quad \frac{n}{n+1}$$

Solution:

The limit has ∞/∞ indeterminate form. Dividing each term of numerator and denominator by n,

$$\Rightarrow \frac{n}{n+1} = \frac{\frac{n}{n}}{\frac{n}{n} + \frac{1}{n}} = \frac{1}{1 + \frac{1}{n}}$$

As n→∞, 1/n→0. Hence,

$$\Rightarrow \lim_{n \to \infty} \quad \frac{n}{n+1} = 1$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to 0} \quad \cos x$$

Solution:

The limit has determinate form. Since cosx is a continuous function, its limit is equal to its value at x=0 i.e. cos0 = 1. Hence,

$$\Rightarrow \lim_{x \to 0} \quad \cos x = 1$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to 0} \quad \log_{0.5} x$$

Solution:

The limit has determinate form. Here, $\log_{0.5} x$ is a continuous function. The base of exponential function is less than 1. As x approaches zero, the function approaches positive infinity (refer its graph).

$$\Rightarrow \lim_{x \to 0} \quad \log_{0.5} x = \infty$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to 0} \quad \frac{1}{x-1}$$

Solution:

Note that we are not testing limit at singularity. The function is determinate form. Hence, limit is :

$$\Rightarrow \lim_{x\to 0} \quad \frac{1}{x-1} = \frac{1}{0-1} = -1$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to\infty} \quad \frac{x^2}{x^2}$$

Solution:

The limit has ∞/∞ indeterminate form. Simplifying,

$$\Rightarrow \lim_{x\to\infty} \quad 1 = 1$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to 0} \quad \frac{\left|x^3\right|}{x}$$

Solution:

The limit has 0/0 indeterminate form. For $x < 0, \quad \left|x^3\right| = -x^3$. Hence,

$$\Rightarrow \frac{\left|x^3\right|}{x} = -\frac{x^3}{x} = -x^2$$

$$\Rightarrow \lim_{x\to 0-} \quad -x^2 = 0$$

For $x > 0, \quad |x^3| = x^3$. Hence,

$$\Rightarrow \frac{|x^3|}{x} = \frac{x^3}{x} = x^2$$

$$\Rightarrow \lim_{x \to 0+} \quad x^2 = 0$$

Clearly, $L_l = L_r = L$. Thus,

$$\Rightarrow \lim_{x \to 0} \quad \frac{|x^3|}{x} = 0$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to 1} \quad \frac{(x^2 - 1)}{|x - 1|}$$

Solution:

The limit has 0/0 indeterminate form. For
$x < 1, \quad |x - 1| = -(x - 1)$. Hence,

$$\frac{(x^2 - 1)}{|x - 1|} = \frac{(x^2 - 1)}{-(x - 1)} = -(x - 1)$$

$$\Rightarrow \lim_{x \to 1} \quad \frac{(x^2 - 1)}{|x - 1|} = \lim_{x \to 1} \quad - (x + 1) = -2$$

For $x > 1, \quad |x - 1| = (x - 1)$. Hence,

$$\frac{(x^2 - 1)}{|x - 1|} = \frac{(x^2 - 1)}{(x - 1)} = (x - 1)$$

$$\Rightarrow \lim_{x \to 1+} \quad \frac{\left(x^2-1\right)}{|x-1|}=\lim_{x \to 1+} \quad (x+1)=2$$

Clearly, $L_l \neq L_r$. Hence, given limit does not exists.

Limits of algebraic functions

Algebraic expressions comprise of polynomials, surds and rational functions. For evaluation of limits of algebraic functions, the main strategy is to work expression such that we get a form which is not indeterminate. Generally, it helps to know “indeterminate form” of expression as it is transformed in each step of evaluation process. The moment we get a determinate form, the limit of the algebraic expression is obtained by plugging limiting value of x in the expression. The approach to transform or change expression depends on whether independent variable approaches finite values or infinity.

The point of limit determines the way we approach evaluation of limit of a function. The treatment of limits involving independent variable tending to infinity is different and as such we need to distinguish these limits from others. Thus, there are two categories of limits being evaluated :

1: Limits of algebraic function when variable tends to finite value.

2: Limits of algebraic function when variable tends to infinite

Limits of algebraic function when variable tends to finite value

In essence, we shall be using following three techniques to determine limit of algebraic expressions when variable is approaching finite value – not infinity. These methods are :

1: Simplification or rationalization (for radical functions)

2: Using standard limit form

3: Canceling linear factors (for rational function)

We should be aware that if given function is in determinate form, then we need not process the expression and obtain limit simply by plugging limiting value of x in the expression. Some problems can be alternatively solved using either of above methods.

Simplification or rationalization (for radical functions

We simplify or rationalize (if surds are involved) and change indeterminate form to determinate form. We need to check indeterminate forms after each simplification and should stop if expression turns determinate. In addition, we may use following results for rationalizing expressions involving surds :

$$\sqrt{a}-\sqrt{b}=\frac{(a-b)}{\left(\sqrt{a}+\sqrt{b}\right)}$$

$$a^{1/3}-b^{1/3}=\frac{(a-b)}{\left(a^{2/3}+a^{1/3}b^{1/3}+b^{2/3}\right)}$$

Example:
Problem : Determine limit

$$\lim_{x\to 1}\frac{\left(\sqrt{x}-1\right)(2x-1)}{2x^2-x-1}$$

Solution : Here, indeterminate form is 0/0. We simplify to change indeterminate form and find limit,

$$\Rightarrow \frac{\left(\sqrt{x}-1\right)(2x-1)}{2x^2-x-1} = \frac{\left(\sqrt{x}-1\right)(2x-1)}{(x-1)(2x+1)} = \frac{(2x-1)}{\left(\sqrt{x}+1\right)(2x+1)}$$

This is determinate form. Plugging “1” for x, we have :

$$\Rightarrow L = \frac{1}{6}$$

Example:
Problem : Determine limit

$$\lim_{x\to 0} \frac{1}{x(8+x)^{\frac{1}{3}}} - \frac{1}{2x}$$

Solution : The indeterminate form is ∞-∞. Simplifying, we have :

$$f(x) = \frac{2-(8+x)^{\frac{1}{3}}}{2x(8+x)^{\frac{1}{3}}}$$

We know that :

$$a^{1/3} - b^{1/3} = \frac{(a-b)}{\left(a^{2/3} + a^{1/3}b^{1/3} + b^{2/3}\right)}$$

Using this identity :

$$\Rightarrow 2-(8+x)^{\frac{1}{3}} = 8^{\frac{1}{3}} - (8+x)^{\frac{1}{3}}$$

$$= \frac{(8-8-x)}{\left(8^{\frac{2}{3}} + 8^{\frac{1}{3}}8^{\frac{1}{3}} + 8^{\frac{2}{3}}\right)}$$

Substituting in the given expression,

$$= \frac{-x}{2x(8+x)^{\frac{1}{3}}\left(8^{\frac{2}{3}} + 8^{\frac{1}{3}}8^{\frac{1}{3}} + 8^{\frac{2}{3}}\right)}$$

$$= \frac{-1}{2(8+x)^{\frac{1}{3}}\left(8^{\frac{2}{3}} + 8^{\frac{1}{3}}8^{\frac{1}{3}} + 8^{\frac{2}{3}}\right)}$$

This is a determinate form. Plugging “0” for x,

$$L = \frac{-1}{2X8^{\frac{1}{3}}\left(8^{\frac{2}{3}} + 8^{\frac{1}{3}}8^{\frac{1}{3}} + 8^{\frac{2}{3}}\right)} = -\frac{1}{48}$$

Example:
Problem : Determine limit :

$$\lim_{x \to 0} \frac{\sqrt{(1-x^2)} - \sqrt{(1-x)}}{x}$$

Solution : Here, indeterminate form is 0/0. We simplify to change indeterminate form and find limit,

$$\frac{\sqrt{(1-x^2)} - \sqrt{(1-x)}}{x} = \frac{\left(1-x^2-1+x\right)}{x\left\{\sqrt{(1-x^2)} + \sqrt{(1-x)}\right\}} = \frac{(1-x)}{\left\{\sqrt{(1-x^2)} + \sqrt{(1-x)}\right\}}$$

This simplified form is not indeterminate. Plugging "0" for "x" :

$$L = \frac{1}{2}$$

Using standard limit form

There is an important algebraic form which is used as standard form. The standard form is (n is rational number) :

$$\lim_{x \to 0} \frac{x^n - a^n}{x-a} = na^{n-1}$$

For rational "n", the expansion of the expression in the limit is given by :

$$\Rightarrow \frac{x^n - a^n}{x-a} = x^{n-1} + ax^{n-2} + a^2x^{n-3} + \ldots\ldots\ldots\ldots + a^{n-1}$$

The expression on the right hand side of the equation is not indeterminate. Thus, limit is obtained simply by plugging "a" for "x" :

$$\Rightarrow L = a^{n-1} + aXa^{n-2} + a^2a^{n-3} + \ldots\ldots\ldots\ldots + a^{n-1}$$
$$\Rightarrow L = na^{n-1}$$

Example:
Problem : Determine limit

$$\lim_{x \to a} \frac{x^{3/2} - a^{3/2}}{x^{1/2} - a^{1/2}}$$

Solution : Here, indeterminate form is 0/0. Substituting $y = x^{1/2} \quad \text{and} \quad b = a^{1/2}$.

$$\frac{x^{3/2} - a^{3/2}}{x^{1/2} - a^{1/2}} = \frac{y^3 - b^3}{y-b}$$

As $\quad x- > a, \quad y- > b$. Using formula,

$$L = 3b^2 = 3\left(a^{1/2}\right)^2 = 3a$$

Example:

Problem : Determine limit

$$\lim_{h \to 0} \frac{(x+h)^{1/n} - x^{1/n}}{h}$$

Solution : Here, indeterminate form is 0/0. Rearranging and using formulae,

$$\lim_{h \to 0} \frac{(x+h)^{1/n} - x^{1/n}}{x+h-x}$$

As $h \to 0$,$x+h \to x$. Using formulae,

$$L = \frac{1}{n} x^{\frac{1}{n}-1}$$

Canceling linear factors (for rational function)

If expression is a rational function, then it is likely that both numerator and denominator become zero at x=a such that given expression has 0/0 indeterminate form. Clearly, then (x-a) is a factor of both numerator and denominator. Canceling common linear factor, we get determinate form. We evaluate limit by plugging limiting value of x in the expression.

Example:
Problem : Determine limit :

$$\lim_{x \to 3} \frac{x^3 - 3x^2 - x + 3}{x^2 - x - 6}$$

Solution : Here, indeterminate form is 0/0. The numerator and the denominator individually tend to 0 as x →3. It means (x-3) is factor of both numerator and denominator. Dividing polynomials (long method or otherwise), we find the quotient. We replace expression in terms of linear factor and quotient :

$$\frac{x^3 - 3x^2 - x + 3}{x^2 - x - 6} = \frac{(x-3)\left(x^2-1\right)}{(x-3)(x+2)}$$

$$= \frac{\left(x^2-1\right)}{(x+2)}$$

This is determinate form. Plugging “3” for “x”, we have :

$$L = \frac{8}{5}$$

Limits of algebraic function when variable tends to infinity

In this case, we are required to estimate behavior of expression when variable approaches very large positive or negative value. The basic idea is to obtain each term in “reciprocal form”. As variable approaches infinity, the reciprocal term approaches zero. There are two methods. However, we need to simplify expression before using either of these methods.

We should also note that these two methods are completely equivalent. We can use either of two methods to evaluate limits. Application of a method is a matter of choice and ease.

1: Dividing each term by highest power of variable

We divide each term of numerator and denominator by x raised to highest positive power in the given expression. Then, we use following limit,

$$x \rightarrow \infty, \frac{c}{x^n} \rightarrow 0; \quad n > 0$$

2: Take out highest power of variable

We take out x raised to highest power from numerator and denominator separately. The highest power variable is considered separately for numerator and denominator. This is unlike previous case in which we use highest power variable of the whole expression. Finally, we evaluate resulting expression using limit rules specified above for the reciprocals and using following additional limits,

$$\text{When} \quad x \rightarrow \infty,$$
$$x^n \rightarrow 0 \quad \text{if} \quad n < 0$$
$$x^n \rightarrow 1 \quad \text{if} \quad n = 0$$
$$x^n \rightarrow \infty \quad \text{if} \quad n > 0$$

We should again emphasize that two methods are essentially equivalent.

Radical and negative variable

Dividing radical by variable poses difficulty when variable represents negative value. Such is the case, when we are evaluating limit in which variable is approaching negative infinity. Clearly, variable has negative value in such cases. We use following rule :

$$\text{If} \quad x < 0, \quad \text{then} \quad x = -\sqrt{(x^2)}$$

In order to understand working of this rule, let us consider a radical :

$$\sqrt{(x^2 + 4x)}$$

We are required to divide the radical by x, when it is known to be negative. Following the fact stated above,

$$\frac{\sqrt{(x^2+4x)}}{x} = -\sqrt{\left(\frac{x^2+4x}{x^2}\right)} = -\sqrt{\left(1 + \frac{4}{x}\right)}$$

Example:
Problem : Determine limit :

$$\lim_{x \to \infty} \frac{x^4 + 2x^3 + 3}{2x^4 - x + 2}$$

Solution : Here, indeterminate form is ∞/∞. Dividing each term by x^4

$$\frac{x^4 + 2x^3 + 3}{2x^4 - x + 2} = \frac{1 + \frac{2}{x} + \frac{3}{x^4}}{2 - \frac{1}{x^3} + \frac{2}{x^4}}$$

This is determinate form. As $x \to \infty, 2/x \to 0, \frac{3}{x^4} \to 0$,Numerator $\to 1$ and as $x \to \infty, \frac{1}{x^3} \to 0, \frac{2}{x^4} \to 0$,Denominator $\to 2$. Hence, limit is :

$$L = \frac{1}{2}$$

Alternatively, we can employ second method to evaluate limit. Taking out x^4

$$\frac{x^4 + 2x^3 + 3}{2x^4 - x + 2} = \frac{x^4\left(1 + \frac{2}{x} + \frac{3}{x^4}\right)}{x^4\left(2 - \frac{1}{x^3} + \frac{2}{x^4}\right)}$$

$$L = \frac{1}{2}$$

Example:

Problem : Determine limit :

$$\lim_{x \to \infty} \frac{a^n + b^n}{a^n - b^n}; \quad a > b > 1$$

Solution : Here, indeterminate form is ∞/∞. By inspection, we see that a/b > 1 and b/a < 1. As we know x→∞,c^x → 0 if c < 1. Hence, we are required to get terms in the form b/a raised to some power. Dividing numerator and denominator by a^n, we have :

$$\frac{a^n + b^n}{a^n - b^n} = \frac{1 + \left(\frac{b}{a}\right)^n}{1 - \left(\frac{b}{a}\right)^n}$$

This is determinate form. As $n \to \infty, (b/a)^n \to 0$. Hence,

$$L = 1$$

Example:

Problem : Determine limit :

$$\lim_{x \to \infty} \frac{1^2}{n^3} + \frac{2^2}{n^3} + \frac{3^2}{n^3} + \ldots\ldots\ldots + \frac{n^2}{n^3}$$

Solution : The indeterminate form is ∞/∞. Writing expression of sum of square of natural numbers, we have :

$$\frac{1^2}{n^3} + \frac{2^2}{n^3} + \frac{3^2}{n^3} + \ldots\ldots\ldots + \frac{n^2}{n^3} = \frac{1^2 + 2^2 + 3^2 + \ldots\ldots + n^2}{n^3}$$

$$= \frac{n(n+1)(2n+1)}{2n^3} = \frac{(n+1)(2n+1)}{2n^2} = \frac{\left(2 + \frac{3}{n} + \frac{1}{n^2}\right)}{2}$$

As $n \to \infty, 3/n, 1/n^2 \to 0$

$$L = \frac{2}{2} = 1$$

Example:
Problem : Determine limit :

$$\lim_{x \to -\infty} \left\{ \sqrt{(x^2+4x)} - \sqrt{(x^2-4x)} \right\}$$

Solution : Here, indeterminate form is ∞-∞. Rationalizing, we have :

$$\sqrt{(x^2+4x)} - \sqrt{(x^2-4x)} = \frac{8x}{\left\{ \sqrt{(x^2+4x)} + \sqrt{(x^2-4x)} \right\}}$$

Dividing by x. Note x is negative. Hence $x = -\sqrt{(x^2)}$

$$= \frac{8}{\left\{ -\sqrt{(1+\frac{4}{x})} - \sqrt{(1-\frac{4}{x})} \right\}}$$

As $x \to -\infty$, $\frac{4}{x} \to 0.$.

$$L = -4$$

Exercises

Exercise:

Problem: Determine limit

$$\lim_{x \to 1} \frac{1}{x^2-1} - \frac{2}{x^4-1}$$

Solution:

Here, indeterminate form is ∞ - ∞. We simplify to change the form of expression from determinate ,

$$\frac{1}{x^2-1} - \frac{2}{x^4-1} = \frac{1}{(x^2-1)} - \frac{2}{(x^2-1)(x^2+1)}$$

$$= \frac{x^2+1-2}{(x^2-1)(x^2+1)} = \frac{x^2-1}{(x^2-1)(x^2+1)}$$

$$= \frac{1}{(x^2+1)}$$

This form is determinate. Plugging “1” for “x”, we have :

$$L = \frac{1}{2}$$

Exercise:

Problem: Determine limit

$$\lim_{x\to\infty} \sqrt{x}\left\{\sqrt{(x+c)} - \sqrt{x}\right\}$$

Solution:

Here, indeterminate form is ∞-∞. Rationalizing, we have :

$$\sqrt{x}\left\{\sqrt{(x+c)} - \sqrt{x}\right\} = \frac{c\sqrt{x}}{\left\{\sqrt{(x+c)} + \sqrt{x}\right\}}$$

Dividing by $\sqrt{x}$,

$$= \frac{c}{\left\{\sqrt{\left(1+\frac{c}{x}\right)} + 1\right\}}$$

As $x \to \infty, c/x \to 0$

$$L = \frac{c}{2}$$

Exercise:

Problem: Determine limit

$$\lim_{x\to\infty}\left\{x - \sqrt{(x^2+x)}\right\}$$

Solution:

Here, indeterminate form is ∞-∞. Rationalizing surd, we have :

$$\left\{x - \sqrt{(x^2+x)}\right\} = \frac{\left\{x - \sqrt{(x^2+x)}\right\}\left\{x + \sqrt{(x^2+x)}\right\}}{\left\{x + \sqrt{(x^2+x)}\right\}}$$

$$= \frac{\left(x^2 - x^2 - x\right)}{\left\{x + \sqrt{(x^2+x)}\right\}} = \frac{(-x)}{\left\{x + \sqrt{(x^2+x)}\right\}}$$

Dividing each of terms by x, we have :

$$= \frac{(-1)}{\left\{1 + \sqrt{\left(1+\frac{1}{x}\right)}\right\}}$$

This is determinate form. As x->∞, 1/x -> 0.

$$L = \frac{-1}{1+1} = -\frac{1}{2}$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to \infty} \left\{ \sqrt{\left(x + \sqrt{x}\right)} - \sqrt{x} \right\}$$

Solution:

Here, indeterminate form is ∞-∞. Rationalizing surds, we have :

$$\left\{ \sqrt{\left(x + \sqrt{x}\right)} - \sqrt{x} \right\} = \frac{\left\{ \sqrt{\left(x + \sqrt{x}\right)} - \sqrt{x} \right\} \left\{ \sqrt{\left(x + \sqrt{x}\right)} + \sqrt{x} \right\}}{\left\{ \sqrt{\left(x + \sqrt{x}\right)} + \sqrt{x} \right\}}$$

$$= \frac{\left(x + \sqrt{x} - x\right)}{\left\{ \sqrt{\left(x + \sqrt{x}\right)} + \sqrt{x} \right\}} = \frac{\sqrt{x}}{\left\{ \sqrt{\left(x + \sqrt{x}\right)} + \sqrt{x} \right\}}$$

Dividing numerator and denominator by $\sqrt{x}$,

$$= \frac{1}{\left\{ \sqrt{\left(1 + \frac{1}{\sqrt{x}}\right)} + 1 \right\}}$$

This is determinate form. As $x \to \infty$, $\frac{1}{x} \to 0$,

$$L = \frac{1}{2}$$

Exercise:

Problem: Determine limit

$$\lim_{x \to a} \frac{x^{5/2} - a^{5/2}}{x^{1/2} - a^{1/2}}$$

Solution:

Here, indeterminate form is 0/0. We put $y = x^{1/2}, b = a^{1/2}; x \to a, y \to b$.

$$\frac{x^{5/2} - a^{5/2}}{x^{1/2} - a^{1/2}} = \frac{x^5 - a^5}{x - a}$$

Using formulae :

$$L = 5b^4$$

$$L = 5\left(a^{1/2}\right)^4$$

$$L = 5a^2$$

Exercise:

Problem: Determine limit

$$\lim_{x\to 1} \frac{1}{x-1} - \frac{2}{x^2-1}$$

Solution:

Here, indeterminate form is ∞ - ∞. We simplify to change indeterminate form and find limit,

$$\frac{1}{x-1} - \frac{2}{x^2-1} = \frac{1}{x-1} - \frac{2}{(x-1)(x+1)}$$

$$= \frac{x+1-2}{(x-1)(x+1)} = \frac{x-1}{(x-1)(x+1)}$$

$$= \frac{1}{(x+1)}$$

It is determinate form. Plugging “1” for “x”, we have :

$$L = \frac{1}{2}$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to\infty} \frac{x^p + x^{p-1} + 1}{x^q + x^{q-1} + 1}; \quad p > 0, q > 0$$

Solution:

Here, indeterminate form is ∞/∞. We divide each term by x^p .

$$\frac{x^p + x^{p-1} + 1}{x^q + x^{q-1} + 1} = \frac{x^p\left(1 + \frac{1}{x} + \frac{1}{x^p}\right)}{x^q\left(1 + \frac{1}{x} + \frac{1}{x^q}\right)}$$

As $x \to \infty$, $\frac{1}{x}$ and $\frac{1}{x^p} \to 0$, and

$$x^{p-q} \to \infty; \quad \text{if} \quad p > q$$

$$x^{p-q} \to 1; \quad \text{if} \quad p = q$$

$$x^{p-q} \to 0; \quad \text{if} \quad p < q$$

Hence,

$$L = \infty; \quad \text{if} \quad p > q$$

$$L = 1; \quad \text{if} \quad p = q$$

$$L = 0; \quad \text{if} \quad p < q$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to 0}\frac{(1+x)^3-1}{3x+2x^2}$$

Solution:

Here, indeterminate form is 0/0. Using standard form,

$$\frac{(1+x)^3-1}{3x+2x^2}=\frac{(1+x)^3-1}{(1+x)-1}X\frac{x}{3x+2x^2}$$

$$=\frac{(1+x)^3-1}{(1+x)-1}X\frac{1}{3+2x}$$

This is determinate form. $x\to 0{,}1+x\to 1$

$$L=\frac{3X1^2}{3}=1$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to 2}\frac{\sqrt{(x-2)}+\sqrt{x}-\sqrt{2}}{\sqrt{(x^2-4)}}$$

Solution:

Here, indeterminate form is 0/0. We simplify to change indeterminate form and find limit,

$$\frac{\sqrt{(x-2)}+\sqrt{x}-\sqrt{2}}{\sqrt{(x^2-4)}}=\frac{\sqrt{(x-2)}}{\sqrt{(x^2-4)}}+\frac{\sqrt{x}-\sqrt{2}}{\sqrt{(x^2-4)}}$$

$$=\frac{1}{\sqrt{(x+2)}}+\frac{\left(\sqrt{x}-\sqrt{2}\right)\left(\sqrt{x}+\sqrt{2}\right)}{\left(\sqrt{x}+\sqrt{2}\right)\sqrt{(x^2-4)}}$$

$$=\frac{1}{\sqrt{(x+2)}}+\frac{\sqrt{(x-2)}}{\left(\sqrt{x}+\sqrt{2}\right)\sqrt{(x+2)}}$$

This is determinate form. Plugging “2” for x, we have :

$$L=\frac{1}{2}$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to 1} \frac{x-1}{\sqrt{(x^2-1)}+\sqrt{(x-1)}}$$

Solution:

Here, indeterminate form is 0/0. Rationalizing surds,

$$\frac{x-1}{\sqrt{(x^2-1)}+\sqrt{(x-1)}} = \frac{\frac{(x-1)}{(x-1)}X(x-1)}{\sqrt{\left\{\frac{(x^2-1)}{(x-1)}X(x-1)\right\}}+\sqrt{\left\{\frac{(x-1)}{(x-1)}X(x-1)\right\}}}$$

Each term limits to na^{n-1} .

This is determinate form.

$$L = \frac{0}{\sqrt{2}+1} = 0$$

Exercise:

Problem: Determine limit

$$\lim_{x \to 64} \frac{x^{1/6}-2}{x^{1/3}-4}$$

Solution:

Here, indeterminate form is 0/0. Using standard formulae, we have :

$$\frac{x^{1/6}-2}{x^{1/3}-4} = \frac{x^{1/6}-64^{1/6}}{x^{1/3}-64^{1/3}}$$

$$= \frac{\frac{\frac{x^{1/6}-64^{1/6}}{x-64}}{x^{1/3}-64^{1/3}}}{x-64}$$

It is determinate form. Evaluating, we have :

$$L = \frac{\frac{1}{6}X64^{1/6-1}}{\frac{1}{3}X64^{1/3-1}}$$

$$L = \frac{1}{4}$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to 1}\frac{\left\{\sqrt{(x^2+8)}-\sqrt{(10-x^2)}\right\}}{\left\{\sqrt{(x^2+3)}-\sqrt{(5-x^2)}\right\}}$$

Solution:

Here, indeterminate form is ∞-∞. Rationalizing surds,

$$\frac{\left\{\sqrt{(x^2+8)}-\sqrt{(10-x^2)}\right\}}{\left\{\sqrt{(x^2+3)}-\sqrt{(5-x^2)}\right\}}=\frac{\left\{\sqrt{(x^2+8)}-\sqrt{(10-x^2)}\right\}\left\{\sqrt{(x^2+8)}+\sqrt{(10-x^2)}\right\}\left\{\sqrt{(x^2+3)}\right.}{\left\{\sqrt{(x^2+8)}+\sqrt{(10-x^2)}\right\}\left\{\sqrt{(x^2+3)}-\sqrt{(5-x^2)}\right\}\left\{\sqrt{(x^2+3)}-\right.}$$

$$=\frac{\left(x^2+8-10+x^2\right)\left\{\sqrt{(x^2+3)}+\sqrt{(5-x^2)}\right\}}{\left(x^2+3-5+x^2\right)\left\{\sqrt{(x^2+8)}+\sqrt{(10-x^2)}\right\}}$$

$$=\frac{\left(2x^2-2\right)\left\{\sqrt{(x^2+3)}+\sqrt{(5-x^2)}\right\}}{\left(2x^2-2\right)\left\{\sqrt{(x^2+8)}+\sqrt{(10-x^2)}\right\}}$$

$$=\frac{\left\{\sqrt{(x^2+3)}+\sqrt{(5-x^2)}\right\}}{\left\{\sqrt{(x^2+8)}+\sqrt{(10-x^2)}\right\}}$$

This is in determinate form. Plugging “1” for “x”, we have :

$$L=\frac{2+2}{3+3}=\frac{2}{3}$$

Exercise:

Problem: Determine limit :

$$\lim_{x\to 1}\frac{x^7-2x_5+1}{x^3-3x_2+2}$$

Solution:

Here, indeterminate form is 0/0. The numerator and denominator tend to 0 as x->1. It means (x-1) is factor of both numerator and denominator. Dividing polynomials (long method or otherwise) and using quotient :

$$\frac{x^7-2x_5+1}{x^3-3x_2+2}=\frac{(x-1)\left(x^6+x^5-x^4-x^3-x^2-x-1\right)}{(x-1)(x^2-2x-2)}$$

$$=\frac{\left(x^6+x^5-x^4-x^3-x^2-x-1\right)}{(x^2-2x-2)}$$

This is determinate form. Plugging “1” by “x”, we know :

$$L=\frac{-3}{-3}=1$$

Exercise:

Problem: Determine limit :

$$\lim_{x \to a} \frac{\sqrt{(a+2x)} - \sqrt{(3x)}}{\sqrt{(3a+x)} - 2\sqrt{x}}; \quad x \neq a$$

Solution:

Indeterminate form is 0/0. We simplify the expression to change indeterminate form and find limit,

$$\frac{\sqrt{(a+2x)} - \sqrt{(3x)}}{\sqrt{(3a+x)} - 2\sqrt{x}} = \frac{\left\{\sqrt{(a+2x)} - \sqrt{(3x)}\right\}\left\{\sqrt{(a+2x)} + \sqrt{(3x)}\right\}\left\{\sqrt{(3a+x)} + 2\sqrt{x}\right\}}{\left\{\sqrt{(a+2x)} + \sqrt{(3x)}\right\}\left\{\sqrt{(3a+x)} - 2\sqrt{x}\right\}\left\{\sqrt{(3a+x)} + 2\sqrt{x}\right\}}$$

$$= \frac{(a+2x-3x)\left\{\sqrt{(3a+x)} + 2\sqrt{x}\right\}}{(3a+x-4x)\left\{\sqrt{(a+2x)} + \sqrt{(3x)}\right\}}$$

$$= \frac{(a-x)\left\{\sqrt{(3a+x)} + 2\sqrt{x}\right\}}{3(a-x)\left\{\sqrt{(a+2x)} + \sqrt{(3x)}\right\}}$$

$$= \frac{\left\{\sqrt{(3a+x)} + 2\sqrt{x}\right\}}{3\left\{\sqrt{(a+2x)} + \sqrt{(3x)}\right\}}$$

This is not in indeterminate form. Plugging “a” for “x”

$$L = \frac{\left\{2\sqrt{a} + 2\sqrt{a}\right\}}{3\left\{\sqrt{(3a)} + \sqrt{(3a)}\right\}} = \frac{\left\{4\sqrt{a}\right\}}{6\sqrt{(3a)}}$$

$$L = \frac{2}{3\sqrt{3}}$$

Continuous function

Idea of "continuity" in the context of function is same as its dictionary meaning. It simply means that function is continuous without any abrupt or sudden change in the value of function. An indicative way to test continuity is to see (graphically or otherwise) that there is no sudden jump in function value in the domain of function. If domain of function is a continuous interval i.e. no points or sub-intervals are excluded from real number set representing domain, then we can draw a continuous function without lifting the drawing instrument i.e. pen, pencil etc. If we have to lift the pen in order to complete the graph of a function in the continuous interval, then function is not continuous.

The feature of continuity is related to function. Therefore, continuity of a function is meaningful in its domain only. It means that we do not need to evaluate continuity in intervals or points where function is not defined. This is an important consideration that helps us to differentiate between “discontinuity” and “undefined values”.

Important concepts

The condition of continuity is expressed in terms of limit and function value. Both of these are required to exist and be equal. We shall learn about these aspects more in detail after having brief overview of these terms.

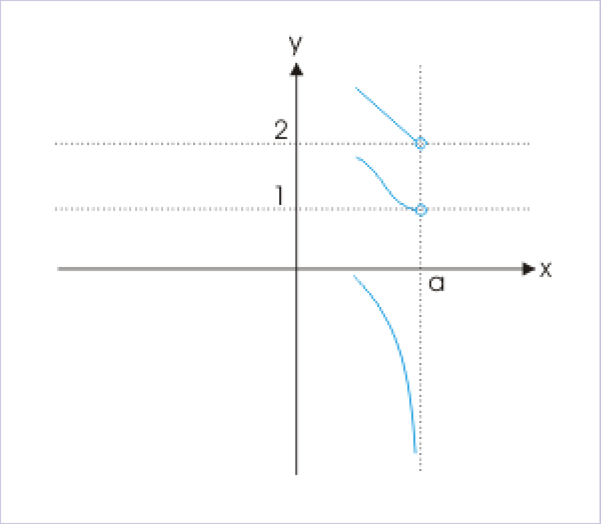
Limit from left

The limit from left means that a function approaches a value (L_l) as x approaches the test point “a” from left such that x is always less than “a” – but not equal to “a”.

$$\lim_{x>a-} f(x) = L_l$$

We show here three illustrations of “limit from left”. The important aspect of these figures is that graph tends to a particular value (infinity is also included). This is done by showing the orientation of graph pointing to limiting value when x is infinitesimally close to test point. Important point to note is that graph does not reach limiting value. Note empty circle at the end of graph, which represents the value of limit not yet occupied by graph. Similarly, asymptotic nature of graph tending to infinity maintains a small distance away from asymptotes, denoting that graph does not reach limiting value.

Limit from left



Limit from left

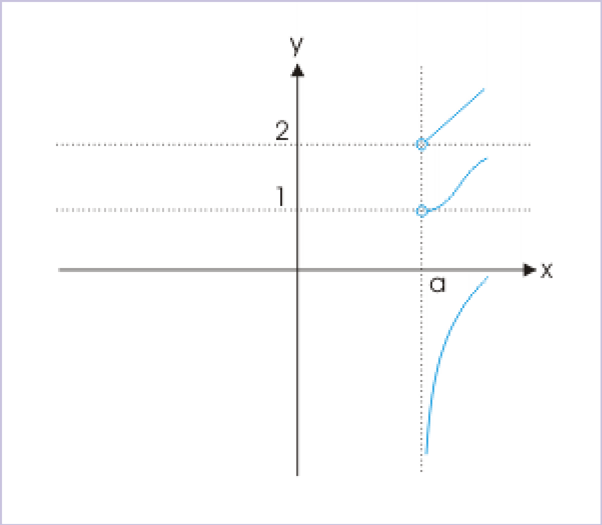
Limit from right

The limit from right means that a function approaches a value (L_r) as x approaches the test point “a” from right such that x is always greater than “a” – but not equal to “a”.

$$\lim_{x>a+} f(x) = L_r$$

We show here three illustrations of “limit from right" as in the earlier case. Important point to note is that graph does not reach limiting value, which represents the value of limit not yet occupied by graph.

Limit from right



Limit from right

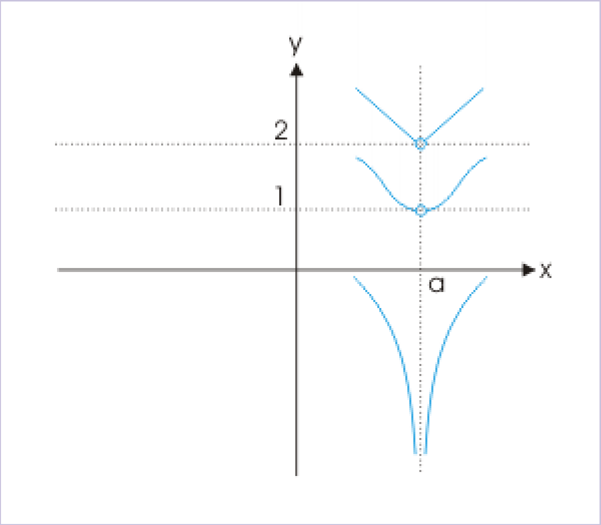
Limit at a point

The limit at a point means that a function approaches a value (L) as x approaches the test point “a” from either side.

$$\lim_{x>a} f(x) = L_l = L_r = L$$

We show here three illustrations of limit at a point. Important point to again note is that graph does not reach limiting value, which represents the value of limit not yet occupied by graph.

Limit at a point



Limit at a point

Function value

Function value is obtained by substituting x values in the function. In case of rational function, we first reduce expression by removing common factors from numerator and denominator.

Continuity at a point

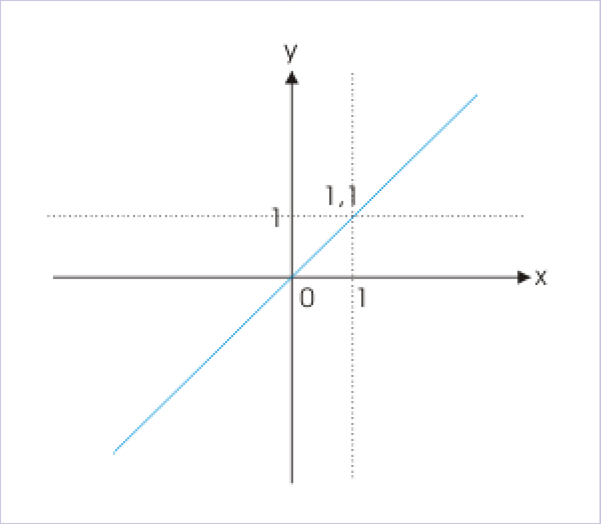
The requirement of continuity is that there should not be abrupt change in function value when there is small change in independent variable. We can enforce this requirement if we can determine x-values in its immediate neighborhood in the domain of function for smallest change in the function values. Mathematically, we say that a function is continuous at a point x=a, if there is small change in function such that $|f(x) - f(a)| < \delta$, then independent variable "x" varies in its immediate neighborhood such that $|x - a| < \in$, where δ and ∈ are arbitrarily chosen small positive numbers.

Condition of continuity is expressed in terms of definition of various limits. Note that limit approaches a value when independent variable comes very close to a point where continuity is being checked. If limit approaches very close to the function value at a point, then it is guaranteed that there exists a value of independent variable in its immediate neighborhood. This fulfills the requirement of continuity as explained in previous paragraph.

For the sake of understanding the requirement of continuity, let us consider identity function, which is known to be continuous in its domain.

$$f(x) = x$$

Identity function



Identity function

Let us consider a test point, x=1. Here, both left and right limit exist and is equal to 1. As such limit of function exists and is equal to 1, which is equal to the function value. As a matter of fact, these observations underline the requirement of continuity at a point. The conditions for continuity at a point in the domain of function are :

1: Limit of function exits at the point.

2: Limit of function is equal to function value at that point.

Mathematically,

$$\lim_{x->a-} f(x) = \lim_{x->a+} f(x) = \lim_{x->a} f(x) = f(a)$$

One important aspect of the requirement is that we test continuity at a finite real value of x, having finite function value. Hence, it is implicitly implied that limit of function should evaluate to a finite function value.

Continuity from left

A function is continuous from left at x=a when left limit exists at x=a and is equal to function value at that point.

$$\lim_{x>a-} f(x) = f(a)$$

Continuity from right

A function is continuous from right at x=a when right limit exists at x=a and is equal to function value at that point.

$$\lim_{x>a+} f(x) = L_r$$

Continuity .vs. limit

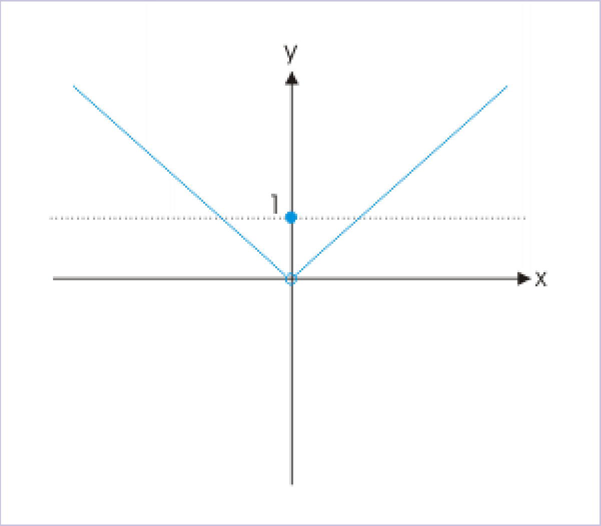
The condition of continuity given above appears to be same as that of limit, which is defined as :

$$\lim_{x->a-} f(x) = \lim_{x->a+} f(x) = \lim_{x->a} f(x) = L$$

However, there is one differentiating aspect. The limit need not evaluate to function value as required for continuity. It means continuity of function has stricter requirement than that of the existence of limit. To understand this point, we consider a variant of modulus function as given here,

`| x ; x>0 f(x) = | 1 ; x=0 | -x ; x<0`

Graph of function



Graph of function

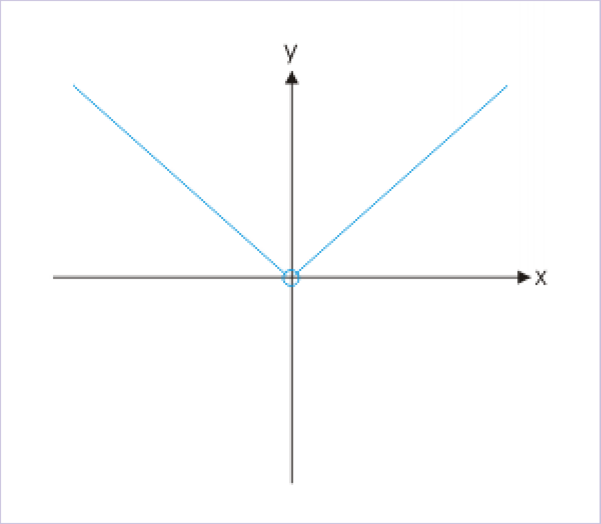
Clearly, limit exists and is equal to zero, but function value is 1 at x=0. Thus, limit is finite, but not equal to function value. As such, given function is not continuous at x=0. The important point to note is that existence of limit or function value at a point is a necessary condition, but not a sufficient condition for continuity. Both the conditions as enumerated should be fulfilled. The concept of continuity, therefore, is a stricter property of a function with respect to limit.

Continuity and exception (singularity)

In order to investigate continuity at singularity point, we consider a function definition which is obtained by modifying modulus function :

`| x; x>0 f(x) = | | -x: x<0`

Graph of function

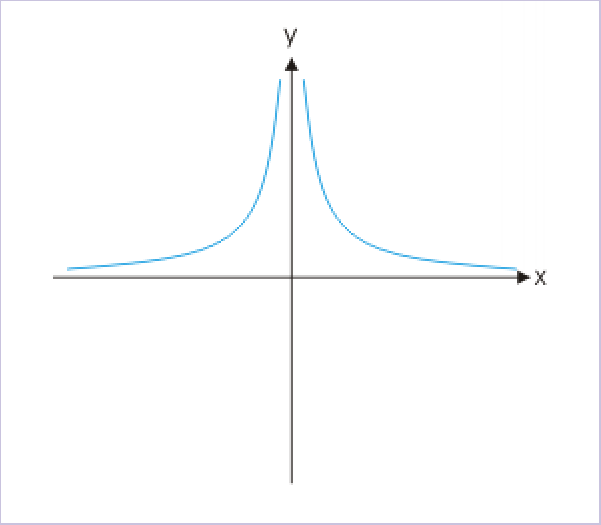


Graph of function

The function is not defined for x=0. Singularity is a point where function is not defined. Thus, this point is singularity for the function which is otherwise defined for all other points on R. They are equal and are equal to a value which is not defined! In accordance with the definition, the limit of function exists at x=0 and is equal to zero. At this point, both left and right limits exist. After all, limit points to a value. Here, it points to a value outside the domain of function. See graph. Existence of limit at x=0, however, has nothing to do with continuity of function at that point as function is not defined at x=0 in the first place. This point is not the part of function definition i.e. its domain and hence continuity or discontinuity is not a concern.

We consider another function. This is modulus of reciprocal condition, f(x) = |1/x|. Using transformation technique, we draw of graph of function as shown :

Graph of function



Graph of function

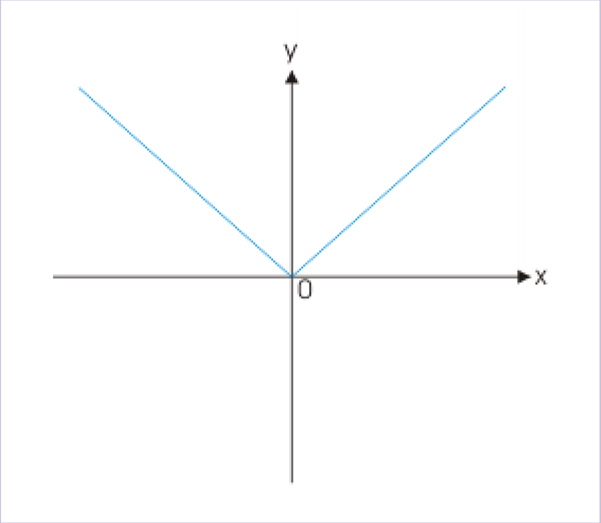
The left and right limits both are positive infinity (read tending to infinity) at x=0. The limit of the function is infinity at this point. But, again point x=0 is not part of function definition. Hence, we say that function is continuous in its domain R-{0}.

Clearly, we need to distinguish between “discontinuous” and “undefined”. Going by two illustrations above, we need to understand that tangent, cotangent, secant and cosecant functions are continuous functions though they appear to be discontinuous graphically. They are not defined at certain values, but then these points are not the part of domain as well. As a matter of fact, rational functions, known to have singularities corresponding to points where denominator is zero, are continuous functions in their domain. For this reason, function such as reciprocal function "1/x" is a continuous function in its domain, which is R – {0}.

Continuity and differentiability

Continuity at a point does not guarantee that function is differentiable at the point. In order to understand this, we now consider the modulus function

itself. Is function continuous at x=0? Is function is differentiable at x=0?
Graph of function



Graph of function

The limit of function is 0, which is finite and is equal to function value. Clearly, function is continuous at this point – thought we can not draw a tangent at this point and as such function is not differentiable at the point. The converse of the assertion, however, is true. If a function is differentiable at a point, then function is continuous at that point. Clearly, differentiability has stricter requirements than that of continuity.

Types of discontinuity

A function is discontinuous if it is not continuous. We can fail the conditions of continuity in many ways. Accordingly, there are following types of discontinuity :

1. Removable discontinuity : Limit of the function exists and is finite, but is not equal to function value. We can remove this type of discontinuity by suitably redefining function value at the test point.

Example:
Problem : Find whether the given function is continuous at x = -2
`| 3x – 2; x ≠ -2 f(x) = | | -4 ; x = -2`
Solution : Here, left and right limits, when x->2, are :

$$L_l = L_r = L = 3X - 2 - 2 = -8$$

Function value at x=-2 is :

$$f(-2) = -4$$

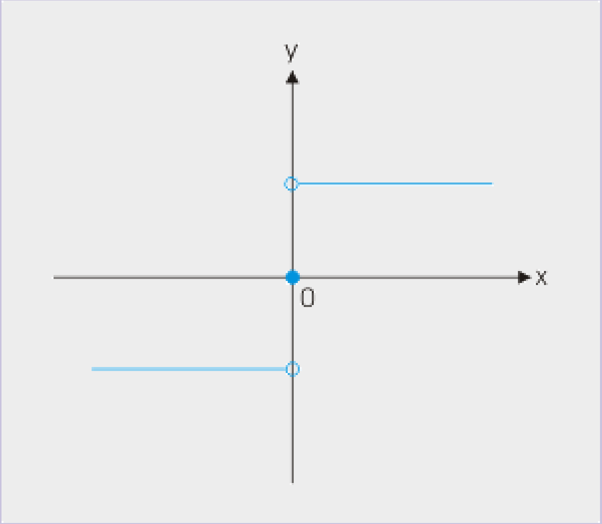
Thus, function is not continuous at x=-2. The discontinuity is removable as we can remove discontinuity by redefining function, at x=-2 as f(x) = -8.
`| 3x – 2; x≠ -2 f(x) = | | -8 ; x = -2`

2. Irremovable or jump discontinuity : This kind of discontinuity arises when left and right limits are not equal. This means limit of function does not exist.

Example:
Problem : Find whether the given function is continuous at x = 0
`| |x|/x; x≠0 f(x) = | | 0 ; x = 0`
Solution : As a matter of fact, this is signum function. For x <0, |x| = -x, Hence, left limit is :

$$\lim_{x>a-} \frac{-x}{x} = -1$$

Graph of function



Graph of function

We see that left limit is not equal to f(0) = 0. We can, therefore, conclude at this stage of analysis itself that function is not continuous at x=0. However, we continue to evaluate right hand limit as well to determine the nature of discontinuity. For x >0, |x| = x. Hence, right limit is :

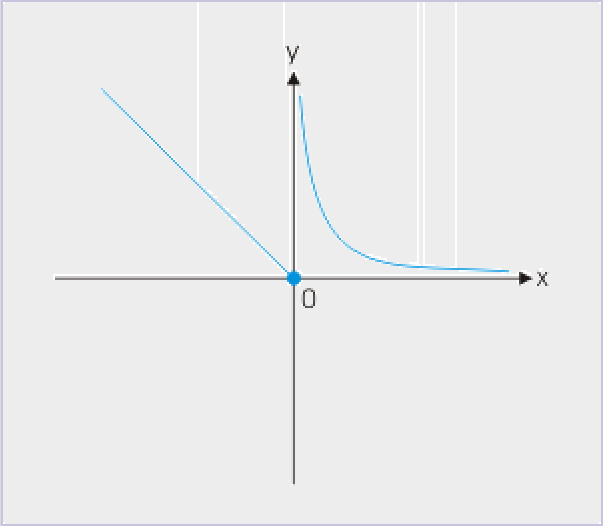
$$\lim_{x>a+} \frac{x}{x} = 1$$

Clearly, $L_l \neq L_r$. The discontinuity is, thus, irremovable or jump type.

3. Essential discontinuity : In this case, at least one of left or right limits does not exist or is infinite. We need to evaluate these conditions in the domain only.

Example:
Problem : Find whether the given function is continuous at x = 0.
`| 1/x; x>0 F(x) = | 0 ; x = 0 | -x; x<0`
Solution : Here, left limit is :

Graph of function



Graph of function

$$\lim_{x > 0-} f(x) = \lim_{x > 0-} x = 0$$

Right limit is :

$$\lim_{x > 0-} f(x) = \lim_{x > 0-} \frac{1}{x} = \infty$$

Since right limit is infinite, the function is discontinuous at x=0.

From these illustrations, it is clear that existence of discontinuity is associated with the manner function is defined. Here, all functions, which are discontinuous at point, are defined in piece-wise manner. On the other hand, basic functions having single definition which we have covered in the course and which are not piece wise defined are continuous functions. We do not intend to generalize these observations, but we can underline that piece - wise definitions indicate possibility of discontinuity.

Further, we note that function value exists and function is defined at the point where function is discontinuous. If there is no function value at a point, then function is not defined at that point and there is no question of continuity or discontinuity.

Continuity in an open interval (a,b)

A function is continuous in an open interval if function is continuous at all points in the interval. This is a simple extension of the concept of continuity at a point. Both left and right limits exist and are equal to function value at all points in the interval. Since end points are not defined, there is always a point on either sides of a given point anywhere in the interval.

Many of the known functions are continuous in open interval. Polynomial, trigonometric, exponential, logarithmic functions etc. are continuous functions in open interval.

Continuity in a closed interval [a,b]

The possibility that there always exist a point around a given point is not there at end points of closed interval. We can not determine left limit at lower end and right limit at upper end of the closed interval. For this reason, we test continuity of function at the closing points from one side only. For a function to be continuous in the closed interval, it should be continuous at all points in the interval and also at the bounding values of closed intereval, [a,b]. Hence,

(i) limit exists at all points in the interval and are equal to function values at those points.

$$\lim_{x>c} f(x) = f(c); a < c < b$$

(ii) right limit exists at x=a and is equal to function value at the lower end of closed interval. .

$$\lim_{x>a+} f(x) = f(a)$$

(iii) left limit exists at x=b and is equal to function value at the upper end of closed interval.

$$\lim_{x>b-} f(x) = f(a)$$

Function operations, compositions of function and continuity

If two functions are continuous at a point or in interval, then function resulting from function operations like addition, subtraction, scalar product, product and quotient are continuous at that point. Further, properly formed function compositions of two or more functions are also continuous.

These properties of continuity are extremely helpful tool for determining continuity of more complicated functions, which are formed from basic functions. Idea is that we are aware of continuity of basic functions. Therefore, continuity of functions formed from these basic functions will also be continuous.

Generally, basic functions are continuous in real numbers set R or its subsets. For example, we know that polynomial functions, sine, cosine, tangent, exponential and logarithmic functions etc are continuous on R. Similarly, a radical function is continuous for non-negative x values. Their composition or the new function will be continuous in the new domain, which is defined in accordance with the rule given here :

- scalar product (multiplication with a constant) : domain remains same
- addition/subtraction/product : domain is intersection of individual domains
- division or quotient : domain is intersection of individual domains minus points for which denominator is zero
- fog or gof : domain is intersection of individual domains

In the nutshell, the function formed from other functions is continuous in new domain as defined above. If we look closely at the definition of continuity here, then "finding interval in which function is continuous" is same as finding "domain" of new function arising from mathematical operations.

Continuous extension of function

Many functions are not defined at singularities. For example, rational functions are not defined for values of x when denominator becomes zero. By including these singular points or exception points in the domain, we can redefine function such that it becomes continuous in the extended domain. This extension of the domain of function such that function remains continuous is known as continuous extension.

Example:

Problem : Obtain the continuous extension of the function given by :

$$f(x) = \frac{x^2 - 1}{x - 1}; \quad x \ne 1$$

Solution : The denominator is zero when x=1. In order that this point is included in the domain of the function, the value of function at this point should be equal to the limit of the function at this point. Now,

$$\lim_{x \to 1} \frac{x^2 - 1}{x - 1} = x + 1 = 2$$

Hence, extended function is :

$$f(x) = \frac{x^2 - 1}{x - 1}; \quad x \ne 1$$

$$= 2; \quad x = 1$$

Combination of continuous and discontinuous functions

We can have combinations of function resulting from function operations or composition, which involves both continuous and discontinuous functions.

We need to know what would be the nature of resulting function? In general, such combinations result in discontinuous function. It is not important to have a generalization here, but such combinations point to the possibility that a function may have discontinuities.

Let us consider two functions f(x) and g(x). Let also assume that f(x) is a continuous function and g(x) is discontinuous function. We, now, consider the operation as :

$$h(x) = f(x) + g(x)$$

Rearranging, we have :

$$\Rightarrow g(x) = h(x) - f(x)$$

In order to test the nature of h(x), let us assume that h(x) is a continuous function. In that case, we know that difference of two continuous functions is a continuous function. As such, g(x) is a continuous function. But, this is contradictory to what we had assumed in the beginning. It means that our supposition that h(x) is continuous function, is wrong. Clearly, function h(x) resulting from addition operation is a discontinuous function. We can extend similar conclusion to subtraction operation as well as subtraction is equivalent to addition operation.

We have already studied few discontinuous functions like greatest integer, least integer and fraction part functions etc. They are discontinuous at integral values. A composition such as y = sin[x] is a discontinuous function. We can analyze these compositions analytically. However, graphical methods are more efficient in this case. We can draw these compositions with the help of transformation of graph by these discontinuous functions. We have dealt this aspect separately in modules titled such as “transformation of graphs by greatest integer function” or “transformation of graphs by fraction part function” etc. For this reason, we shall not discuss this topic any further here in this module.

Examples

Example:

Problem : Find whether the given function is continuous at x = -2

`| g(x); x ≠ 1 f(x) = | | 2 ; x = 1`

where,

$$g(x) = \frac{x^3 - 1}{x^2 - 1}$$

Solution : In order to factorize cubic expression in the numerator, we guess that x=1 is one real root. We can see that expression turns zero for x=0. Hence, we conclude that (x-1) is one of the factor of cubic expression. Now, for x ≠ 1,

$$f(x) = \frac{x^3 - 1}{x^2 - 1} = \frac{(x-1)\left(x^2 + x + 1\right)}{(x-1)} = \frac{\left(x^2 + x + 1\right)}{(x+1)}$$

The reduced expression is not an indeterminate form. Hence, left and right limits, when x->1, are :

$$L_l = L_r = L = \frac{3}{2}$$

Here, $f(1) = 2$. We, therefore, conclude that function is not continuous at x=1. This is a removable discontinuity as we can remove this discontinuity by redefining function at x=1 as f(x) = 3/2.

Example:

Problem : Find whether the given function is continuous at x = 0

`| x sin(1/x), x ≠ 0 f(x) = | | 0; x = 0`

Solution : For x ≠ 0;

$$\lim_{x>0-} x \sin\left(\frac{1}{x}\right) = 0$$

Note that as x-->0, 1/x-->infinity and sin(1/x) --> a finite value in [-1,1]. Thus, function tends to become 0 X finite value, which is equal to zero. Similar is case for right limit. Hence,

$$L_l = L_r = L = f(0) = 0$$

Thus, function is continuous at x = 0.

Exercises

Exercise:

Problem:

If the function given below is continuous in its domain, then determine value of k.

$$f(x) = \frac{x^3 + x^2 - 3x + 3}{(x-1)^2}; \quad x \neq 1$$

$$= k; \quad x = 1$$

Solution:

For x ≠1, the function is :

$$f(x) = \frac{x^3 + x^2 - 3x + 3}{(x-1)^2}$$

We guess here that one of the roots of numerator is 1. It is true as numerator is zero for x=1. Thus, (x-1) is a factor of cubic expression in the numerator.

$$\Rightarrow f(x) = \frac{(x-1)\left(x^2 + 2x - 3\right)}{(x-1)^2} == \frac{(x-1)(x-1)(x+3)}{(x-1)^2}$$

$$\Rightarrow f(x) = x + 3$$

Clearly,

$$\Rightarrow \lim_{x \to 1} f(x) = \lim_{x \to 1} x + 3 = 4$$

For function to be continuous, this limit should be equal to value of function at x=1. Hence,

$$\Rightarrow f(1) = k = 4$$

$$\Rightarrow k = 4$$

Exercise:

Problem: Determine if the given function is continuous.

$$f(x) = \frac{x}{\sin x}; \quad x \neq 0$$

$$= 0; \quad x = 0$$

Solution:

For x ≠ 0, the function is of standard form whose limit is 1 when x approaches 0. Thus, limit is not equal to function value at x=0. Clearly, function is discontinuous at x=0. It is a removable discontinuity.

Exercise:

Problem: Determine continuity of function given at x= 0 :

$$f(x) = \frac{e^{1/x}}{1 + e^{1/x}}; x \neq 0$$

$$= 0; \quad x = 0$$

Solution:

In order to evaluate limit at x=0, we determine left and right limits at x=0. We see here that as x approaches zero 1/x approaches negative infinity from left, whereas it approaches to positive infinity from the right. Using these facts, left hand limit is :

$$\Rightarrow \lim_{x\to 0-} \frac{e^{-\infty}}{1+e^{-\infty}} = \frac{0}{1+0} = 0$$

Now right limit is :

$$\Rightarrow \lim_{x\to 0+} \frac{e^{1/x}}{1+e^{1/x}} \quad \left[\frac{\infty}{\infty} \quad \text{form}\right]$$

Dividing by $e^{1/x}$, we have :

$$\Rightarrow \lim_{x\to 0+} \frac{1}{\frac{1}{e^{1/x}}+1} = \frac{1}{\frac{1}{e^{\infty}}+1} = 1$$

Clearly, $L_l \neq L_r$. Hence, function is discontinuous at x=0.

Exercise:

Problem: Determine continuity of the function given by :

$$f(x) = x; \quad \text{x is rational}$$

$$= 2 - x; \quad \text{x is irrational}$$

Solution:

In order to determine continuity, we consider set of rational and irrational numbers separately. Let us first work with rational set. We determine limit when as x approaches any real value point “a” in real domain. Note that x approaches real number “a” assuming only rational number in its set. We say that x approaches "a" through rational numbers. Then,

$$\Rightarrow \lim_{x\to a} f(x) = \lim_{x\to a} x = a$$

Now, we work with irrational set. Here, x approaches real number “a” assuming only irrational numbers in the domain. Then,

$$\Rightarrow \lim_{x\to a} f(x) = \lim_{x\to a} 2 - x = 2 - a$$

Let us consider that function is continuous at x=a. In that case,

$$L_l = L_r$$

$$\Rightarrow a = 2 - a$$

$$\Rightarrow a = 1$$

We have taken any arbitrary point x=a and we find that function is continuous only for a single value – not an interval or union of intervals. Thus, we conclude given function is continuous only at one point and discontinuous at all other points.